INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Training users of a flexible decision support system: The effects of training content and user ability

Rodriguez, Suzanne Marie, Ph.D.

Case Western Reserve University, 1991

Copyright ©1991 by Rodriguez, Susanne Marie. All rights reserved.
TRAINING USERS OF A
FLEXIBLE DECISION SUPPORT SYSTEM:
THE EFFECTS OF TRAINING CONTENT AND USER ABILITY

by
SUZANNE MARIE RODRIGUEZ

Submitted in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy

Thesis Advisor: Dov Te’eni

Department of Management Information and Decision Systems
CASE WESTERN RESERVE UNIVERSITY
May 17, 1991
CASE WESTERN RESERVE UNIVERSITY

GRADUATE STUDIES

We hereby approve the thesis of

Suzanne Marie Rodriguez

candidate for the  Ph.D. in Management

degree.*

Signed: ____________________________

(Chairman)

[Signature]

Date  April 9, 1991

*We also certify that written approval has
been obtained for any proprietary material
contained therein.
TRAINING USERS OF A
FLEXIBLE DECISION SUPPORT SYSTEM:
THE EFFECTS OF TRAINING CONTENT AND USER ABILITY

Abstract

by

SUZANNE MARIE RODRIGUEZ

Two forms of training are contrasted in this study: operations training and strategy training. Operations training provides information concerning how to operate the features of the Decision Support System (DSS). Operations-trained subjects have sole responsibility for the development of a decision strategy. Strategy training provides information concerning the operation of DSS features in the context of a suggested decision strategy. Strategy-trained subjects need not develop their own strategy for using the DSS. The dimension of user ability selected for study was cognitive restructuring ability (CRA) as measured by the Group Embedded Figures Test. CRA is used in the strategy formulation activities of selecting, organizing and reorganizing information. CRA was expected to interact with the form of training in determining performance because the training treatments provided different levels of support to the
strategy formulation process. Forty-two MBA students participated in a laboratory experiment based on a job sequencing task. Computer logs, questionnaire responses and concurrent verbal protocols were analyzed. Strategy training resulted in higher performance, regardless of CRA. For operations-trained subjects, high CRA individuals performed better than low CRA individuals. The relationship with CRA was reversed for strategy-trained subjects. A number of underlying strategy differences between groups were also documented. The findings suggest that strategy training can be used as a means of responding to individual differences in the user population.
To my parents and grandparents, who taught me all that I know of dedication, sacrifice, and the value of hard work; and to my husband, Rolando, who reminds me always of the value of laughter.
# TABLE OF CONTENTS

## Chapter 1. Introduction

1.1 Existing Responses to User Differences .................................. 2  
1.2 Efficiency versus Effectiveness ........................................... 3  
1.3 User Adaptation .......................................................... 7

## Chapter 2. Literature Review

2.1 The Problem Posed by a System to its User ............................ 14  
2.2 DSS Design Features and Problem Structure ........................... 18  
2.3 Individual Differences and Problem Structuring ....................... 22  
   2.3.1 Cognitive Restructuring Ability ......................... 23  
   2.3.2 Findings Regarding CRA in an MIS/DSS Context .............. 37  
   2.3.3 Conclusions Regarding CRA in an MIS/DSS Context .......... 47  
2.4 User Training and Problem Structure ................................. 51  
   2.4.1 Instructional Psychology .................................. 52  
   2.4.2 End-User Training ......................................... 60  
   2.4.3 Conclusions Regarding End-User Training .................... 75  

Tables/Figures for Chapter .............................................. 81
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4.5.5</td>
<td>Perceived Adequacy of Domain Knowledge</td>
<td>189</td>
</tr>
<tr>
<td>5.4.6</td>
<td>Maintenance and Generalization of Training</td>
<td>190</td>
</tr>
<tr>
<td>Tables/Figures for Chapter</td>
<td></td>
<td>192</td>
</tr>
</tbody>
</table>

**Chapter 6: Data Analysis**  
200-289

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Descriptive Statistics</td>
<td>200</td>
</tr>
<tr>
<td>6.2</td>
<td>Overview of Analysis Procedures</td>
<td>201</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Statistical Analysis Procedures</td>
<td>202</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Protocol Analysis Procedures</td>
<td>204</td>
</tr>
<tr>
<td>6.3</td>
<td>Cost Performance</td>
<td>215</td>
</tr>
<tr>
<td>6.4</td>
<td>Time Performance</td>
<td>219</td>
</tr>
<tr>
<td>6.5</td>
<td>The Strategy Formulation Process</td>
<td>223</td>
</tr>
<tr>
<td>6.5.1</td>
<td>Information Use</td>
<td>223</td>
</tr>
<tr>
<td>6.5.1.1</td>
<td>Expected Differences by Training Treatment</td>
<td>225</td>
</tr>
<tr>
<td>6.5.1.2</td>
<td>Information Use within Training Treatments</td>
<td>227</td>
</tr>
<tr>
<td>6.5.2</td>
<td>User Strategies</td>
<td>230</td>
</tr>
<tr>
<td>6.5.2.1</td>
<td>Concurrent Verbal Protocols</td>
<td>230</td>
</tr>
<tr>
<td>6.5.2.2</td>
<td>User Logs</td>
<td>238</td>
</tr>
<tr>
<td>6.5.2.2.1</td>
<td>Net Effect Operator Use</td>
<td>238</td>
</tr>
<tr>
<td>6.5.2.2.2</td>
<td>Best Net Effect Operator Use</td>
<td>240</td>
</tr>
<tr>
<td>6.5.2.2.3</td>
<td>Calculations</td>
<td>240</td>
</tr>
</tbody>
</table>
6.6 The Strategy Execution Process .......... 242
  6.6.1 Concurrent Verbal Protocols .......... 242
  6.6.2 User Logs ............................ 243
6.7 User Perceptions .......................... 244
  6.7.1 Confidence ............................. 245
  6.7.2 Perceived Usefulness
      of Information ........................... 246
      6.7.2.1 Perceived Importance .......... 246
      6.7.2.2 Perceived Usableness .......... 248
  6.7.3 Perceived Difficulty .................. 248
  6.7.4 Overall Satisfaction
      with Strategy ............................ 249
  6.7.5 Perceived Adequacy of
      Domain Knowledge ........................ 251
6.8 Maintenance & Generalization
      of Training .............................. 252
  6.8.1 Maintenance ............................ 253
  6.8.2 Generalization ........................ 255
6.9 Summary of the Findings .................. 259
Table/Figures for Chapter .................... 262

Chapter 7. Conclusion ........................ 290-300

  7.1 Contributions and Limitations ........... 290
  7.2 Discussion ............................... 294
LIST OF FIGURES

Figure 1. Examples of ordinal and disordinal interactions .......................... 84

Figure 2. Research framework ............................................. 100

Figure 3. Structure of the decision support system screens .......................... 135

Figure 4. Schedule Comparison Screen .................................... 136

Figure 5. Modification Screen ............................................ 137

Figure 6. Analysis Screen ................................................ 138

Figure 7. Changeover Costs Screen ...................................... 139

Figure 8. High/Low Values Screen ...................................... 140

Figure 9. Flowchart of Stage I of the suggested strategy ...................... 141

Figure 10. Flowchart of Stage II of suggested strategy ....................... 142

Figure 11. Changeover alternatives considered in Stage II ..................... 143

Figure 12. Flowchart of Stage III of suggested strategy ..................... 144

Figure 13. "Up" alternatives considered in Stage III ......................... 145
Figure 14. "Down" alternatives considered in Stage III .......................... 146

Figure 15. Confidence items screen  ......................... 193

Figure 16. Expected cost performance outcomes by treatment  ................. 194

Figure 17. Distribution of GEFT scores across treatments ..................... 263

Figure 18. Group cost performance means across job sets ...................... 267

Figure 19. Graphs of cost performance regression lines ....................... 272

Figure 20. Time performance means across job sets .......................... 274
# LIST OF TABLES

Table 1. List of Acronyms .............................. 81

Table 2. Summary of CRA Studies in an MIS/DSS Context ........... 82-83

Table 3. Summary of MIS/DSS Training Studies ..................... 85-87

Table 4. Evaluations of the Experimental Tasks Used in Previous GEFT Research ................... 132-133

Table 5. Job Set Parameters ............................... 134

Table 6. Subject Assignments to Treatment-by-Ability Groups ........ 192

Table 7. Expected Differences in Information Item Use by Training Treatment .................. 195

Table 8. Expected Differences in Information Item Use Within the Operations Training Treatment ... 196

Table 9. Descriptions of High and Low Analytic Decision Processes ........ 197

Table 10. Summary of Predictions ......................... 198-199

Table 11. Descriptive Statistics for the Sample .................... 262
Table 25. Results of the Statement-by-
Statement Protocol Analysis ............ 282

Table 26. Summary of Protocol Subjects’
Methods for Incorporating
Net Effect Information
into the Decision ................... 283-284

Table 27. Regression Analysis Results for
Net Effect Operator Use ............... 285

Table 28. Analysis of Contrasts within the
Net Effect Operator Use Data ........... 286

Table 29. Analysis of Contrasts within the
Best Net Effect Operator Use Data .... 287

Table 30. Regression Analysis Results for
Calculations Performed' ............... 288

Table 31. Average Overall Satisfaction
Strategy ............................... 289
CHAPTER 1. INTRODUCTION

This study is concerned with enhancing decision-making effectiveness through an understanding of the role of individual differences during the formulation of decision strategies. This work contributes to the discussion of how to respond to individual differences during the design and implementation of management information systems (MIS) and decision support systems (DSS). Because individual differences amongst the multiple users of a system increase "the risk of implementation failure or misuse" (Moore & Chang, 1983, p. 203), the determination of appropriate responses to user differences is an important research problem.

This chapter briefly summarizes existing responses to individual differences, suggests a primary limitation of those responses, and questions the assumptions underlying those responses. The second chapter presents a review of literature concerning human-computer interaction, the individual difference selected for study, and the design of training. The third chapter introduces the research framework and research questions. The fourth chapter describes the system and task environment developed for this study. The fifth chapter explains the experimental methods
employed to test hypotheses. The sixth chapter summarizes the analysis procedures and experimental findings. Finally, the seventh chapter presents a discussion of the findings, the limitations of the study, and implications for research and practice.

1.1 EXISTING RESPONSES TO USER DIFFERENCES

Two major responses to individual differences have been advocated (Benbasat & Taylor, 1978). The first response necessitates discovering what different types of people need and want in different problem-solving contexts and tailoring systems to individuals. What individuals of various types need has been determined largely through observing performance in the presence of different system design features. A contingency table of staggering proportions would be required to guide system design. Huber (1983) has gone so far as to suggest that this approach will not lead to operational design guidelines.

The second response alleviates the difficulty associated with predicting what users will need or want: Provide a flexible enough system and virtually anything a user needs or wants will be possible. In comparison to the first response, this response appears
both feasible and immediately applicable (within the limits of current technology).

In both of the existing responses, the main idea seems to be not to interfere with the user’s natural decision process. The assumption is that users know best how to structure their own interactions with a system, and it is either not possible or not desirable to intervene. In both approaches, individual differences are addressed during system design, as opposed to system implementation.

1.2 EFFICIENCY VERSUS EFFECTIVENESS

The two responses outlined above can be evaluated in terms of the objectives they are likely to achieve. One common objective for computerized systems is to improve the efficiency of task performance, users’ "ability to accomplish a job with a minimum of time and effort" (Webster’s, 1989). The efficiency of task performance can be enhanced by assisting users in carrying out the steps of their decision strategy, a process which will henceforth be referred to as the strategy execution process.

Both of the responses outlined above attempt to facilitate the strategy execution process without intentionally influencing the nature of the decision
strategy being executed. Flexible designs or tailored designs avoid inefficiencies in strategy execution which result from a mismatch between user needs and wants and system features. Both approaches are consistent with the goal of improving the efficiency of decision making processes. Average efficiency is maximized by supporting the efficient execution of many different decision processes.

However, it may be insufficient to improve decision making efficiency without also achieving the objective of improved decision making effectiveness, particularly in the case of DSS. Decision making effectiveness calls attention to the quality of the decisions made with a system rather than the resources consumed in making those decisions. The quality of a decision is determined mainly by the quality of the underlying decision strategy. To improve decision effectiveness, then, requires attention to the process by which users arrive at their decision strategies, a process which will henceforth be referred to as the strategy formulation process.

Neither of the above responses to individual differences addresses performance differences arising from the quality of decision processes created during strategy formulation. By neglecting the strategy
formulation stage, they fail to consider what "is frequently the most difficult part of solving a problem" (Silver, 1989, p. 18). Yet the goal of improved decision effectiveness cannot be achieved without attending to the nature of users' decision processes.

Strategy formulation is also the stage in which individual differences combine to determine needs and wants. Without considering the formulation stage in some depth, it is impossible to understand the true import of individual differences. Increasingly, attention is being directed to the performance ramifications of individual differences in the repertoire of strategies and strategy formulation skills (Brown, Bransford, Ferrara & Campione, 1983; Sternberg, 1985).

As indicated in the opening sentence of this work, the alternative response to individual differences presented here is intended to achieve the objective of improved decision-making effectiveness. Hence, this study focuses on the strategy formulation process which determines the nature of decision strategies. Training will be used in order to affect the course of the strategy formulation process. In contrast to the two responses discussed above, individual differences will
be addressed during implementation instead of during system design.

The individual difference measure selected for this study, the Group Embedded Figures Test (Witkin, Oltman, Ruskin & Karp, 1971), was chosen primarily because it measures one class of strategy formulation skill. Another reason for selecting the Group Embedded Figures Test (GEFT) was the quantity of existing research, including MIS/DSS studies. The arguments presented in the following pages, however, are intended to transcend any specific individual difference. All individual differences can be understood in terms of their impact on the strategy formulation process. The sum total of what we are determines the process we adopt in response to each situation. The GEFT was chosen merely as a convenient vehicle to demonstrate the comparatively greater role of individual differences during strategy formulation, as opposed to strategy execution.

The departure point for this work is to question a basic assumption implicit in the existing responses to individual differences: Multiple users cannot all make effective use of the same, limited set of design features or will be extremely dissatisfied in doing so. From this viewpoint, changes in users' decision
strategies brought about by the system are considered undesirable and/or difficult to achieve. Rather than simply accepting this proposition, I suggest that we consider more thoroughly the conditions under which it is likely to hold true.

1.3 USER ADAPTATION

Adaptation implies a qualitative shift in the nature of the strategies users employ. Thus, adaptation is accomplished through the strategy formulation process which controls the nature of user strategies. Improvements in decision making effectiveness also require qualitative changes in the nature of decision strategies employed. No improvement in decision making effectiveness can occur without user adaptation. Hence, if we desire enhanced decision effectiveness, we must accept user adaptation as a prerequisite.

When a user is encouraged to employ a decision process different from the "natural" process employed by that user, the desirability of the change depends on the comparative merits of the two processes and how well the user can execute them. Changes in user strategies which result in more effective decisions are certainly desirable. In the interest of improved
decision effectiveness, adaptation on the part of all but the most successful performers is necessary and desirable.

It is important to stress that variability in decision strategies translates into variability in performance. In our haste to emphasize that no single approach or skill profile is better than another for the wide range of tasks confronting a manager, we must remember that one approach or skill profile can be better than another with respect to specific problems, subproblems or stages of decision making. The value of each individual can be affirmed without condemning changes in the "natural" course of specific decision processes.

The driving force behind qualitative change in decision strategies is learning (Keen & Gambino, 1983), so we can consider the difficulty of adapting in terms of the difficulty of learning what an appropriate strategy is. Users must learn how to solve problems with a system. Just as we discuss the difficulty of solving problems in terms of the structure of the problem, the difficulty of learning can be described as a function of the structure of the learning situation.

When referring to the structure of a learning situation, I wish to emphasize factors in addition to
the objective structure of an underlying decision to be made or problem to be solved with a system. These other factors include individual characteristics, system design attributes, and the nature of training provided to users. All of these factors work together to determine the difficulty of learning how to use a computerized tool to solve a problem. The latter two factors, system design attributes and the nature of training, are clearly within the control of system builders. Hence, the difficulty of user adaptation depends partly on the control we exercise over the structure of the learning situation.

Keen and Gambino refer to users who adapt well during prototyping as "good users" (1983). They recognize that "only skilled users can learn how to use a tool in new ways" (Keen & Gambino, 1983, p. 147). What makes a good user? First, there are characteristics of users (e.g., domain knowledge) which make the same situation more structured for some users than others. Second, there are characteristics of users (e.g., motivation) which make some users more willing to structure a situation through learning. Finally, there are characteristics of users (e.g., the skill measured by the GEFT) which make some users better able to structure a situation. In this study, I
consider explicitly only the skill measured by the GEFT.

System designs can contribute structure to a situation by "hinting" at what must be learned to adapt to the situation successfully. Sage suggests that helping the user adapt is "a principal task of a well-designed information system" (1981, p. 659). "Decision channeling" (Stabell, 1983) is a means of directing the course of learning in the hope that a desired set of discoveries will be made by the user. "Strong designs" attempt to structure the learning situation by directing the course of learning, while "weak designs" do not intentionally direct the course of learning (Moore & Chang, 1983). In this study, the difficulty of adaptation is not alleviated through the system design, which is intentionally flexible and "weak."

User training is the most obvious way in which we can direct the course of learning. However, user training can also be "strong" or "weak" with respect to its impact upon the structure of a learning situation. This study contrasts a "strong" form of training, strategy training, with a "weak" form of training, operations training.

As I stated earlier, the structure of the learning situation determines the difficulty of adapting, and
without adaptation there can be no gains in effectiveness. We should not underestimate the difficulty of learning in an ill-structured situation. Without explicit guidance provided through training or system help features, the user learns mainly through an idiosyncratic series of experiences. As will be discussed in the next chapter, this type of learning situation favors high ability users. In addition, Brehmer (1980) has suggested a number of ways in which experience is not the best teacher, regardless of the capabilities of the pupil. "The problems suggest that managers alone and unaided are not likely to be able to systematically improve their decision making" (Stabell, 1983, p. 231). By exercising control over the structure of the learning situation, however, system designers and implementers can establish conditions conducive to learning and improvements in decision making.

The experiment reported here was designed to demonstrate the importance of strategy formulation skill in using a flexible DSS. The experiment also seeks to demonstrate that the relationship between formulation skill and performance is moderated by the content of training. In a larger sense, the experiment was designed to illustrate that individual differences
can be addressed through training with resulting increases in average decision performance. While a training response to individual differences has limitations of its own (as discussed in the concluding chapter), this study suggests that it is a feasible means of enhancing decision-making effectiveness.
CHAPTER 2. LITERATURE REVIEW

Four areas of research are included in this review: the psychology of human-computer interaction, the influence of DSS design attributes on users, the role of users' cognitive restructuring ability, and user training. As described in the introductory chapter, all of these areas must be considered in determining the structure of the learning situation facing users. The structure of the learning situation, which is partially within the control of system designers and implementers, determines the likelihood that users will adapt successfully in the form of effective decision strategies.

The first section of this chapter examines the human-computer interaction problem which must be solved by a system's users, since the skill with which users solve the interaction problem ultimately determines their decision-making performance. The second and third sections discuss the impact of design features and individual differences on the ease with which the interaction problem is solved. The fourth section summarizes how user training has addressed the system-use problem. User training is an important means of controlling the difficulty experienced by users in
adapting to a system. The key findings presented in this chapter lead to the research framework, research questions and hypotheses stated in Chapter 3.

For the reader’s convenience, the acronyms used in this chapter and subsequent chapters are summarized in Table 1. Throughout the text, tables and figures are provided at the end of each chapter in the order in which they are cited in the chapter.

2.1 THE PROBLEM POSED BY A SYSTEM TO ITS USER

The user facing a specific computerized system for the first time is attempting to understand the device itself and the way in which the device should enter into the solution of a problem. Kieras and Polson state that a "new user’s initial attempts to work with a device are under the control of problem-solving processes" (1985, p. 378). The user must develop a device representation and task representation in order to solve the problem. The device representation consists of knowledge concerning the system itself, while the task representation combines device-independent and device-dependent knowledge required to make a decision using the system (Kieras & Polson, 1985).
Kieras & Polson (1985) suggest that the user's task representation or strategy can be described by the GOMS model (Card, Moran & Newell, 1983). The user has a set of goals and subgoals which describe a plan for completing the task, a set of methods which are procedures for achieving the subgoals, a set of selection rules which match the available methods to characteristics of the task, and knowledge of the elementary operations required by each method. In essence, the user must acquire a mental model of the task which hierarchically relates goals, subgoals, methods and elementary operations. The user's strategy consists of combinations of elementary operations bounded by conditions for which the strategy is appropriate. The strategy is most commonly represented as a set of production rules (e.g., Card et al., 1983; Greeno, 1978).

There is no single mechanism through which users acquire mental models. Bostrom, Olfman & Sein (1990, p. 103) suggest at least three "mapping processes"
through which users acquire understanding, separately or in combination:

1) **Mapping via usage**: Users can acquire a mental model of the system merely through using it.

2) **Mapping via analogy**: Users can acquire a mental model of a new system by drawing analogies from similar systems that are familiar to them.

3) **Mapping via training**: Users can acquire a mental model of the system through training.

User characteristics affect all three mapping processes, while system interface characteristics are particularly important during the mapping via usage process (Bostrom et al., 1990). The rest of this section and the next will consider the mapping via usage process, while later sections will also consider mapping via analogy and mapping via training.

The difficulty of devising an adequate strategy for incorporating a system into a specific decision (i.e., of acquiring an appropriate set of GOMS) can be
expressed in terms of the degree to which the problem is structured (Simon & Hayes, 1976):

Not all authors have defined the distinction [between structured and ill-structured problems] in exactly the same way, but a common theme running through all of the definitions is that problems are ill structured to the extent that the problem solver himself must contribute toward their definition.

From this description it is clear that novel tasks, for example, are less structured than familiar tasks (Carlson, 1983; Moore & Chang, 1983). It is also clear that the structure of a problem is subjective, so that a problem changes as the "perceptions of the individual involved change" (Hurst, Ness, Gambino & Johnson, 1983, p. 124).

The designer does not create a system in ignorance of the challenge it poses to its future users. The system image (Norman, 1983) reflects the designer’s understanding of how the system is to be used (Stabell, 1983, p. 224):

How the user should (is expected to) perceive and understand the system--its uses and limitations--in the context of the user’s particular decision situation is also an important aspect of the interface architecture.
The interplay between a system's design and users' strategies has long been recognized (Ackoff, 1967; Churchman, 1969; Payne, 1982; Kottemann & Remus, 1989). The next section briefly considers the role of system interface characteristics in the mapping via usage process.

2.2 DSS DESIGN FEATURES AND PROBLEM STRUCTURE

Silver (1989) has focused attention on some important attributes of systems which describe how strongly the system attempts to shape perceptions or, in other words, how much structure the system contributes to the human-computer interaction. One dimension concerns the overall flexibility of a system design. This dimension, which Silver calls system restrictiveness, is defined as follows (Silver, 1989, p. 9):

System Restrictiveness: the degree to which and the manner in which a Decision Support System limits its users' decision-making processes to a subset of all possible processes.

Systems can restrict by excluding information or functions, constraining the order in which operations can be performed, or limiting user control over parameters and inputs. Because cost-benefit tradeoffs
in strategy execution affect strategy formulation (Payne, 1982; Jarvenpaa, 1989), it is impossible in practice to make a clean distinction between restrictiveness which will affect strategy formulation and restrictiveness which will affect only strategy execution. This work is concerned with the role of restrictiveness in influencing strategy formulation.

The more restrictive a system is, the more strongly it suggests the structure of an appropriate process. The forces to which the user is adapting are more apparent. In contrast, flexibility places the responsibility of structuring the interaction squarely on the shoulders of the user. Hence, the problem of determining how to use a system productively is more difficult (ill-structured) for a flexible system than for a roughly equivalent restrictive system.

In an experimental setting, Dos Santos (1982) showed that user performance may be higher in more structured learning situations: The restrictive version of a decision support system was associated with higher problem understanding than the more flexible version. The flexible version of the DSS allowed users to explore a simulated business environment by altering the values of a number of variables and observing the effects. The restrictive DSS constrained the number of
variables which could be manipulated over time, initially allowing only one variable to be manipulated. The restrictive DSS enforced a strategy of isolating the effects of one variable before considering variables in combination. This is a strategy which psychologists know well as the "separation of variables schema" (Inhelder & Piaget, 1958). Users performed better with the more structured system, and failed to structure the interaction on their own. The junior and senior management students in the DoSantos experiment evidently had quite a bit in common with far younger subjects who also experience difficulty in applying the separation of variables schema (e.g., Linn, 1978).

Another system attribute which alters the balance of structuring responsibility is decisional guidance (Silver, 1989, p. 17):

**Decisional Guidance**: the degree to which and the manner in which a Decision Support System guides its users in constructing and executing decision-making processes, by assisting them in choosing and using its operators.

Silver further classifies decisional guidance according to whether it assists in formulating a strategy versus executing it, or whether it is informational versus suggestive in nature. Both decisional guidance and restrictiveness contribute structure to the system-use
problem by increasing the extent to which the system contributes to the interaction, thereby assisting users in adapting.

In the discussion of problem structure, I emphasized its subjective nature. Similarly, we would expect that the extent to which a system design contributes structure to the learning situation is also in the eyes of the beholder. Silver (1988b) has demonstrated that system restrictiveness is not an absolute concept. He conducted an experiment in which subjects used three different DSS to prepare a ranked list of five cities in which they would like to live. The selections were made from a database of information pertaining to ninety-seven cities. From an expert, objective standpoint, the systems could be strictly ranked in terms of restrictiveness on the basis of the decision-making processes supported by each. User assessments of restrictiveness, however, differed from the objective ranking. The results were explained in terms of user-desired processes and user-accessible processes which differed from the available processes. The importance of a consideration of individual differences was again confirmed.
2.3 INDIVIDUAL DIFFERENCES AND PROBLEM STRUCTURING

The need to consider individual differences in designing computer systems is an established part of the MIS/DSS literature (Mason & Mitroff, 1973; Zmud, 1979; Sage, 1981; Reneau & Grabski, 1987). The types of individual differences thought to affect system use include demographic, personality and cognitive style variables (Benbasat & Taylor, 1982). In recognition of the lack of any existing operational design guidelines which can account for the endless variety of differences in users, flexible designs are often recommended (Benbasat & Taylor, 1978; Sprague & Carlson, 1982; Huber, 1983). Robey suggests that individual differences research "has provided much of the rationale for a flexible DSS" (1983, p. 580). However, flexibility is yet another design attribute which interacts with individual differences in determining performance.

For example, the relationship between user domain experience and flexibility is recognized (Alter, 1981; Paller, 1983). Essentially, a more experienced user is better able to take advantage of a flexible system than a novice because the decision-making process is more fully structured or programmed. Similarly, the expert
is in less need of decisional guidance. In terms of the earlier discussion, the expert has a more complete initial task representation and/or more extensive knowledge facilitating the formulation of an appropriate task representation (Ericsson & Simon, 1984).

Even assuming that all users have equivalent experience, however, flexibility and guidance could still have differential effects on users due to differences in certain, key abilities. Especially for novel tasks where the device-independent portion of the task representation is being created in conjunction with the device-dependent portion, an individual's skill in performing the necessary analytical and structuring activities will determine the quality of the resulting strategy. In such unstructured situations, flexible DSS place great demands on structuring skills. A measure of structuring skill, then, would be useful in understanding the nature of performance with flexible DSS.

2.3.1 COGNITIVE RESTRUCTURING ABILITY

One instance of the type of skill described above is called cognitive restructuring ability (Witkin & Goodenough, 1981). Analysis and structuring are viewed
as complementary aspects of the same tendency towards dealing with a field in an active manner. Cognitive restructuring skills enable the individual to perform three related activities (Witkin & Goodenough, 1981, p. 51):

(1) breaking up an organized field so that its parts are experienced as discrete from background;

(2) providing organization to a field that lacks it; or

(3) imposing a different organization on a field from the one suggested by its inherent organization.

Frank (1983) suggests that cognitive restructuring ability (CRA) represents greater flexibility of information processing.

The process of mapping via usage involves the activities facilitated by CRA. A strategy for system use is the product of an individual’s initial understanding, the meaning discovered through analysis and the meaning assigned by imposing structure, the "refined thinking" that results in a "well developed" problem representation (Ramaprasad, 1987). CRA is important in formulating strategies for using systems and in interpreting the cues provided by systems because the process of decomposing stimuli is important (Einhorn & Hogarth, 1982, p. 23). Those individuals
who possess strong cognitive restructuring skills should be better able to isolate important information items from the context of a screen containing many items of information, organize system outputs and capabilities in terms of a problem-solving approach, and work around limitations in system outputs and capabilities by mentally reorganizing the field.

A confusion in nomenclature has arisen over the course of approximately four decades of research and theory building concerning CRA. CRA is related to a cognitive style called field dependence-independence, but is not equivalent to it (Witkin & Goodenough, 1981, pp. 56-61; Linn & Kyllonen, 1981). Field dependence-independence is a higher level construct referring to the tendency of individuals to function autonomously. According to the theory of field dependence-independence, field dependent individuals tend to develop skills which further an interpersonal orientation, while field independent individuals tend to develop skills, such as cognitive restructuring, which further an autonomous, problem-solving orientation. Field dependence-independence qualifies as a style dimension since neither of these orientations is superior across all situations.
CRA, on the other hand, is unipolar because it is better to have it than not to have it. Individuals with strong cognitive restructuring skills are expected to be at an advantage with respect to many problem-solving situations. The unipolar nature of CRA is in direct contradiction to the philosophy of cognitive styles research that "emphasizes an approach rather than an ability" (Huber, 1983, p. 567).

This work is based on arguments concerning an ability dimension (CRA) rather than implications of the broader field-dependence-independence cognitive style, and I do not consider it cognitive style research. While I am not alone in this belief (Pracht & Courtney, 1988), other researchers have applied the cognitive style label to CRA even when it is explicitly acknowledged as an ability (Benbasat & Dexter, 1979). Perhaps the reluctance in the information systems literature to present tests of CRA as ability tests is linked to a recognition that "ability/disability" terminology can be threatening to the individual being tested (Benbasat & Taylor, 1978; Witkin, Moore, Goodenough & Cox, 1977).

Confusion concerning measurement instruments also pervades research related to field dependence-independence and CRA. Arbuthnot (1972) concluded that
many contradictory or confusing patterns of research results could be attributed to the use of a variety of instruments which are not equivalent. Witkin and Goodenough (1981) suggested that the Embedded Figures Test (Witkin et al., 1971) should be used to measure CRA. The test requires that an individual trace a simple figure located in a distracting, complex field. On the basis of this perceptual disembedding task, subjects can be classified on a continuum from low to high skill (henceforth referred to as low and high analytic). Although the test is based on a perceptual task, extensive evidence indicates that the same skill is manifested in functioning with symbolic representations when thinking and problem solving require cognitive restructuring (Witkin & Goodenough, 1981). Witkin & Goodenough (1981, p. 60) suggested that a battery of tests, in addition to the Embedded Figures Test, should be used in assessing the broader dimension of field dependence-independence.
Because the Embedded Figures Test (EFT) assesses cognitive ability, it has certain advantages over style instruments (Pracht & Courtney, 1988, p. 604):

Cognitive ability, unlike cognitive style, does not vary dramatically from problem to problem, unless the subject develops new abilities during the problem-solving process or in the interval between problems. Thus, findings based on cognitive ability may remain more stable and be easier to replicate than those based on cognitive style.

CRA has been found to stabilize in young adulthood, with high test-retest correlations (Witkin et al., 1971, p. 5).

Ericsson & Simon (1984) have also suggested that assessment of cognitive abilities is more likely to succeed than assessment of style. They point out several differences between style and ability tests in the processes underlying the response to test items. Correct responses to ability tests cannot be generated by guessing or first impressions, the test items can fairly accurately represent the "real" situation, and the items do not require that subjects remember or anticipate behaviors (Ericsson & Simon, 1984, p. 43).

The ability measured by the EFT is regarded as one of the specific abilities that contribute to overall intelligence (Witkin & Goodenough, 1981, p. 61). Not
surprisingly, EFT scores do correlate strongly with "analytic" subtests of standard intelligence tests. Tests of CRA have been associated with fluid (as opposed to crystallized) intelligence (Cronbach & Snow, 1977, p. 382; Witkin & Goodenough, 1981; Linn & Kyllonen, 1981). This relationship is not surprising in view of the obvious conceptual similarities between the two: discussions of fluid intelligence also refer to processes of decontextualization or analysis, organization, and reorganization (Snow & Lohman, 1984). Linn & Kyllonen (1981, p. 270) have suggested a number of characteristics that tests of CRA have in common, including unfamiliar stimulus material and generation of novel representations. Fluid intelligence is also viewed as most relevant to situations demanding flexible adaptation to novel environments (Snow & Lohman, 1984, p. 360). For purposes of this research effort, it is not necessary to enter into a debate concerning the distinction (if any) between CRA and fluid ability. The concepts are clearly related, and Witkin did not deny the relationship.

Because CRA is only one class of abilities, one cannot say that high analytics are superior in general intelligence (Witkin et al., 1971, p. 7). However, the three cognitive restructuring activities identified at
the beginning of this section are important in a variety of contexts. Up until this point I have emphasized the relevance of CRA to the mapping via usage process, but the same activities required in problem solving can be required for learning from instruction (mapping via training). For example, "frequently in learning, the material to be learned lacks clear inherent structure, creating the requirement that the learner himself provide organization as an aid to learning" (Witkin, Moore, Goodenough & Cox, 1977). Because the activities are important to both mapping processes, I have organized the discussion in the remainder of this section in terms of the activities involved rather than attempting to apply the results separately to one mapping process or another.

A number of studies demonstrate the superior ability of high analytcs in separating individual items, particularly relevant ones, from background information. For example, two studies in an educational setting used instruments similar to the EFT to illustrate the link between CRA and the specific activity of separating relevant from irrelevant information. Annis (1979) tested college students' memory of a written text. She found that high
analytics were superior in answering questions of high importance to the meaning of a text, but that low analytics did as well as high analytics on questions of low importance.

The second study examined memory of lecture material under both immediate and delayed testing conditions (Kardash, Lukowski & Bentmann, 1988). An immediate test can provide cues concerning information relevance to guide studying prior to a delayed test. The authors showed that a test immediately following the lecture provided sufficient knowledge of a criterion that low analytics performed as well as high analytics on a delayed test. Without an immediate test to provide cues concerning information relevance, high analytics performed better on a delayed test.

The ability to selectively process information has also been related to the automatization of a simple skill (Jolly & Reardon, 1985). The authors hypothesized that "automaticity requires an ability to isolate essential elements of the behavior to be automatized so that they can be utilized" (Jolly & Reardon, 1985, p. 303). High analytic subjects were faster in forming an automatized sequence. Furthermore, incidental memory for information differed by level of CRA. High analytics remembered less task-
irrelevant information than low analytics. High analytics also remembered increasingly more task-relevant information when forced to cope with varying degrees of task interruption, whereas low analytics showed an opposite tendency.

The results reported by Kiewra & Frank (1988) can be interpreted as illustrating the role of CRA in selecting relevant information and/or in reorganizing information. When college students were tested on the material presented during a lecture, high analytics were better at answering test questions which evaluated higher order knowledge (application, analysis, synthesis and problem solving). Perhaps, as in Annis (1979), high analytics were more likely to remember the general principles necessary to answer difficult questions. Alternatively, high analytics may have been better at reorganizing what they knew in order to apply it. High analytics have been found to be superior in studies directly concerning transfer-of-training (Goodenough, 1976), suggesting that CRA also impacts the process of mapping via analogy (i.e., recognizing similarities between problems encountered during training and new problems).

Frank (1983) investigated one of the implications of reorganizing ability for memory. He found that the
cues required to trigger recall for high analytics are less specific than the cues required for low analytics. Subjects did not differ in the amount remembered under conditions of free recall or when the cue presented during learning was presented during recall. However, high analytics outperformed low analytics when the cue presented during learning was different from (though related to) the cue presented during recall.

Witkin & Goodenough (1981, pp. 27-28) report the superiority of high analytics on two other types of tasks, functional fixity and set-breaking problems, which can also be understood in terms of reorganizing information. Functional fixity problems require that an item (e.g., a bottle stopper) being used to perform one function (seal a bottle) be perceived as capable of another function (serve as a wedge). Set-breaking learning sequences first encourage, then make inappropriate, a specific solution approach. Unless information is reorganized, the subject will continue to apply an inappropriate strategy. The reorganization in both tasks would appear to require a shift in the cues which are considered salient (Greeno, 1978, p. 47). High analytics have been found to sample more
fully from the available cues and to shift more quickly when cue salience is altered (Goodenough, 1976; Frank, 1983).

The organizing ability of high analytics has also been studied in a number of settings. High analytics are more likely to impose structure on materials ranging from Rorschach inkblots to social studies concepts (Witkin et al., 1977). Frank (1984) found that high analytic college students were more likely to employ an outline format in their notes, and that low analytic students performed better when provided with an outline.

By now the reader should have gained some appreciation of the manner in which CRA manifests itself in learning and problem solving through the activities of selection, organization and reorganization. While the evidence of performance differences is considerable, there is still some debate concerning the origin of performance differences. The issue is whether high and low analytics are employing different strategies, or executing the same strategies with different levels of success. Evidence exists in both directions, although surprisingly little research has been devoted to the origins of performance differences.
The origin of performance differences is important in determining the feasibility of overcoming those differences. Evidence of strategy differences associated with CRA indicates that performance differences can be addressed through training (keeping in mind the role of CRA in learning). Any performance differences related to strategy execution would be unaffected by training in appropriate strategies (except as a function of practice).

In the areas of learning and memory, Goodenough's (1976) review of the literature favored the position that CRA determines how individuals learn rather than how much they learn (strategy rather than capacity). According to Goodenough, high analytics are more likely to adopt a hypothesis-testing, "participant" role in concept-attainment, while low analytics assume a passive, "spectator" approach. In keeping with their active approach, high analytics may also be more likely to spontaneously organize information during free recall.

In support of strategy differences, a recent CRA study explicitly examining strategy use (Clark & Roof, 1988) did find differences in the use of global versus analytic strategies. However, the study could not explain all performance differences in terms of the
strategy dimensions examined. Perhaps efficiency
differences linked to strategy automatization might
account for much of the remaining variation, as well as
providing a link between the formulation-differences
and execution-differences viewpoints.

A review by Davis & Frank (1979) acknowledges that
process differences probably underly high analytics'
more efficient memories and greater ability to conduct
combinatorial analysis. Those process differences,
however, determine the likely success of the individual
in terms of how much is learned or remembered. There
is also more of a suggestion in their work that
capacity/efficiency differences unrelated to process
may account for performance differences.

In support of a capacity perspective, one study
found performance differences even after attempting to
hold process constant (Reardon, Jolly, McKinney &
Forducey, 1982). All subjects were forced to take an
active role in learning, but high analytics still
outperformed low analytics on a recall test of
incidental material. The difficulty, of course, is in
determining whether the strategies employed by high and
low analytics were equivalent in all respects.
Additionally, low analytics acting in an unfamiliar,
active learning role may have been engaged in more
effortful processing than high analytics acting in a familiar learning role. The availability of attentional resources might explain the greater memory of high analytics, in which case the difference would be expected to diminish over time.

In summary, this section has discussed the nature of CRA and presented findings illustrating its importance in problem solving and learning. When cognitive restructuring activity is demanded by the situation, performance differences linked to CRA are commonplace. In the absence of such demands (e.g., highly structured material), the link is broken. Strategy differences associated with CRA suggest the possibility of improving performance through training of appropriate strategies.

2.3.2 FINDINGS REGARDING CRA IN AN MIS/DSS CONTEXT

The studies reviewed in this section directly support the relevance of CRA to human-computer interaction. This section includes only those studies which found a relationship between CRA (as measured by either the EFT or the GEFT) and user performance. Because CRA is not a cognitive style (as explained above), I have omitted: (1) a review of cognitive styles literature beyond those studies using the EFT,
and (2) a detailed response to critics of cognitive styles research. In addition to the explanations presented earlier, this decision is justified by the lack of straightforward relationships between CRA and several of the stylistic variables considered in the information systems literature. For example, no consistent relationships have been found between CRA and: locus of control, extraversion/introversion (Witkin & Goodenough, 1981, p. 48), the Learning Style Inventory (Highhouse & Doverspike, 1987), or the Myers-Briggs Type Indicator (Schmidt & McCutcheon, 1988).

In accordance with the terminology preferred in information systems studies, individuals possessing strong cognitive restructuring skills will be referred to as "high analytics." Low CRA individuals will be termed "low analytics." None of the reasoning I present in this document is intended to apply without modification to classifications based on other test instruments using similar terminology (eg., the analytic-heuristic instrument employed by Dickson, Senn & Chervany [1977]).

Two early studies in the management literature examined the effects of various report formats on the performance of high or low analytic subjects (Lusk, 1979; Lusk & Kersnick, 1979). The five formats
investigated ranged from a tabular presentation of raw data to a highly "transformed" graphical presentation. The simple tabular report was perceived by students as least complex, while the three graphical reports were seen as more complex than tabular reports. Subjects were provided with one of the five formats and required to answer twenty questions. Four questions could be answered directly from the report provided, and the other sixteen questions required that the subjects manipulate the data provided to obtain the answers.

As the authors had predicted, high analytics performed better than low analytics in terms of the number of questions answered correctly. There was no evidence of differential performance peaking: both high and low analytics performed better with the less complex report formats. The most interesting aspect of the study from my perspective was the training implication suggested by Lusk & Kersnick (1979, p. 797):

Since the performance of high analytic individuals is sometimes greater but never less than the performance of low analytic individuals for particular tasks, one may wish to train low analytic individuals to become analytic...training sessions may aid individuals in developing their operational heuristics...
The training described by Lusk & Kersnick would change individuals' behavior with respect to specific reports without attempting to alter their underlying analytic ability.

In contrast to Lusk (1979) and Lusk & Kersnick (1979), the other studies reviewed below found both performance differences and evidence of differential performance peaking or differential information preferences. For example, Benbasat, Dexter & Masulis (1981) found differences between low and high analytics in terms of report requests. Subjects could choose either a tabular history report or a graphical report to aid them in making an optimal price/quantity decision. The authors felt that the graphical report was more relevant to the solution of the problem, but that low analytics would be less able to make use of information presented in the graphical format.

Not only did high analytics perform better in terms of successful task completion, but they also made greater use of the graphical report. "In the analysis of report requests over time it was observed that low analytics were the only ones who neither decreased their requests for history reports nor increased their requests for graphical reports as the game progressed" (Benbasat et al., 1981, p. 760). This result
demonstrated that low analytics possessed different format preferences than high analytics. The question of differential performance peaking could not be addressed given the design of the study.

A second study (Benbasat & Dexter, 1985) also investigated graphical versus tabular format with the added consideration of color. In comparison to Benbasat et al. (1981), neither format was considered more relevant to the experimental task (allocating an advertising budget across three territories). Subjects were provided with one of the two report formats, either in multi-color or mono-color.

The results showed an overall advantage for high analytics, and an interaction of color with CRA. The worst performers in terms of potential profit achieved were low analytics with mono-color reports. With multi-color reports, low analytics performed as well as high analytics. High analytics were unaffected by the presence of color. While there were no performance effects linked to report format, low analytics did rate tabular reports as more understandable than graphical reports. There was no difference in ratings for high analytics.

Benbasat & Dexter (1979) compared two methods of providing information: database inquiry for day-by-day
historical information (events approach) versus structured/aggregate reports (value approach). Subjects used systems incorporating one of the two methods in making inventory/production decisions. Differences related to CRA were found for profit performance, time performance, report request behavior and self-reported decision approach.

There was an interaction of CRA and system type similar in pattern to the interaction found in Benbasat & Dexter (1985). High analytics performed equally well with either type of system. Low analytics performed poorly with the structured/aggregate report and made more report requests than high analytics. Low analytics performed as well as high analytics and made the same number of report requests when provided with the database inquiry capability. High analytics also reported "a more stable, fixed strategy approach compared to low analytics" (Benbasat & Dexter, 1979, p. 740).

A subsequent study by Benbasat & Dexter (1982) showed that low analytics could perform as well as high analytics with the structured/aggregate report if they were provided with an additional decision aid. The decision aid the authors provided allowed subjects to explore various scenarios before being committed to a
decision. While both high and low analytics had better performance with the decision aid, the effect was more pronounced for the low analytics.

A study by Pracht and Courtney (1988) reported different findings on the effect of a decision aid. They provided a graphical, interactive, problem-structuring tool aimed at improving decision-environment (problem-structure) understanding. The tool allowed users to create and revise diagrams depicting the relationships in a simulated business environment. Without the tool, high and low analytics performed equally on decision-environment comprehension tests. The tool improved performance for the high analytics, but not for the low analytics.

Pracht & Courtney (1988) suggested that the failure of the tool to improve the performance of low analytics was due to their lower spatial abilities and corresponding inability to develop an adequate visual representation. They stated that the graphical nature of their decision aid could partially account for the difference between their results and the results of the Benbasat & Dexter study (1982). The difference in results was also attributed to the use of a less structured problem environment and the examination of the problem formulation stage rather than the problem
solving stage. (These stages correspond closely, but not exactly, to what I have been calling strategy formulation and strategy execution. As I understand their meaning, problem formulation is the process of developing a problem representation. Strategy formulation requires not only arriving at a problem representation, but also of developing procedures for operating upon it.)

I agree that it is possible to "account for the apparent discrepancy by considering the kind of assistance required by the different groups compared to the kind of help offered by the decision aid" (Pracht & Courtney, 1988, p. 616). However, I do not believe that the differences identified by Pracht & Courtney account for the results. If anything, the less structured problem environment should have provided greater opportunities to assist low analytics. Although Benbasat & Dexter (1982) observed problem solving behavior, Pracht & Courtney themselves wrote that "the structure used to solve the problem determines the final solution" (1988, p. 599). In other words, Benbasat & Dexter's results also depended on the problem structuring stage.

Despite the authors' interpretation, the Pracht & Courtney study (1988) did not reveal whether
differences in spatial ability accounted for the low analytics' failure to benefit from the aid. In order to attribute the results to differences in spatial abilities we would need to know if low analytics were less proficient in representing and understanding graphically presented relationships. If given a set of relationships described verbally, could low analytics depict them graphically with the tool provided? If provided with a graphical depiction of the environment, could low analytics make use of the information in understanding the problem? These questions are not answered by the study.

I would suggest that the discrepancy between the two studies (Pracht & Courtney, 1988; Benbasat & Dexter, 1982) can be explained without attributing the results to any of the three factors identified by Pracht & Courtney. There are two possible reasons for the low analytics' failure to benefit from the Pracht & Courtney decision aid which could apply either separately or in conjunction. First, the tool assists mainly in remembering and depicting relationships. This is of value only after the relationship has been identified, but the difficulty experienced by a low analytic in identifying the relationship in the first
place (decontextualization) is not reduced by a memory aid.

Second, each set of revisions made to the visual representation can be viewed as a refinement of the subject's hypotheses regarding the decision environment. In this sense, the tool is more compatible with a hypothesis-testing approach to learning about the environment. As discussed in the previous section, research concerning concept-attainment and CRA suggests that high analytics use such an approach, but low analytics do not (Goodenough, 1976).

Both of these explanations separate the problem structuring process from the related process of graphically depicting the structure so obtained. The low analytics needed help with problem structuring, but they received help only with depicting problem structure. High analytics were able to benefit from the memory aid because they were able to discover a structure to depict.

In comparison, the Benbasat & Dexter (1982) decision aid reduced the demands on problem structuring abilities. The decision aid largely eliminated the requirement that subjects thoroughly understand and be able to predict relationships in the decision
environment. Subjects could test the impact of various decisions without predicting them or understanding why they worked. As I see it, the Benbasat & Dexter decision aid helped low analytics because they did not have to depend on weak problem structuring abilities. The Pracht & Courtney aid (1988) did not help low analytics because it did not address their difficulty with problem structuring.

2.3.3 CONCLUSIONS REGARDING CRA IN AN MIS/DSS CONTEXT

The studies reviewed in the previous section are summarized in Table 2. The findings summarized above can be used to support the following conclusions:

Conclusion 1. If a performance difference exists for an analytic, problem-solving task, high analytics outperform low analytics.

All of the studies reviewed support this conclusion. Differences in system features (eg. decision aids or presentation format) can sometimes increase the profit performance of low analytics to the level of high analytics (Benbasat & Dexter, 1985; Benbasat & Dexter, 1979; Benbasat & Dexter, 1982). The same studies found that the performance of high analytics was less sensitive to differences in system features.
Conclusion 2. Past research regarding CRA has focused on design, as opposed to training, implications.

The research reviewed above is aimed at developing a set of MIS/DSS design guidelines. For various reasons, however, this design orientation has been severely criticized (Huber, 1983; Ramaprasad, 1987). I agree with Ramaprasad (1987) and Stabell (1983, p. 257) that operational guidelines must be based on specific rather than general influences on information processing. The user's strategy ultimately determines the value of any design feature, but CRA is too general to predict the specific nature of the model that will be developed via usage. The user's level of CRA can only suggest whether the individual is likely to encounter difficulties in the mapping via usage process. Knowledge of CRA does not predict the details of the mapping.

The design orientation in previous CRA research may be related to the confusion between style and ability mentioned earlier. The change in perspective from style to ability has a subtle but fundamental influence on the manner in which one attempts to respond to differences in CRA. When one assumes that the differences between the strategies of high and low analytics are in the manner of goal attainment rather
than the extent of goal attainment, then the focus becomes one of facilitating whatever it is the user should decide to do. There is no "problem" in the mapping via usage process which needs to be addressed.

Given an ability perspective where it is assumed that high analytics possess a skill which facilitates the development of better strategies, the focus becomes one of how to enable low analytics to use strategies more like those of high analytics. Once the strategy is developed, the role of cognitive restructuring ability in strategy execution would appear minor. In relatively few situations involving spatial tasks, such as the use of multi-line graphs, would perceptual disembedding ability appear to be directly linked to strategy execution (eg. Benbasat & Dexter, 1985).

In situations involving symbolic material, the extent to which cognitive restructuring is required would seem to depend on the subjective extent to which an organizing structure or the salience of certain symbols is perceived by the individual. In other words, symbolic embeddedness would appear to be in the eyes of the beholder just as problem structuredness is in the eyes of the beholder. But the strategy itself provides an organizing structure which dictates the separation of item from field. How much symbolic disembedding
occurs during strategy execution once the relationships and important items of information have been determined? In other words, the primary concern from an ability perspective is one of facilitating strategy development rather than strategy execution.

The literature presented in this section suggests that low analytics are at a disadvantage when it comes to models created via usage. We know from the first section of this chapter, however, that strategies can also be acquired through mapping via training. Training represents another possible direction for CRA research which has not yet been explored in the MIS/DSS literature. As mentioned earlier, Lusk & Kersnick (1979) have suggested that training might be used to improve the performance of low analytics by helping them to act like high analytics. Huber (1983) has also identified training/coaching as one area in which the study of individual differences might prove useful.

Conclusion 3. The strategies actually employed by users of differing CRA have not received direct attention.

Just as in the majority of psychological studies discussed in Section 2.3.1, the studies reviewed in the previous section measured attributes of the decision process without revealing the nature of the decision
process. The studies provide little information as to why and how aspects of design affected user performance. Without such an understanding, it is impossible to generalize to other systems and task contexts. Todd & Benbasat (1987) indicate that this weakness is shared by DSS research in general and suggest that process tracing methods could contribute greatly to our knowledge concerning human-computer interaction.

2.4 USER TRAINING AND PROBLEM STRUCTURE

Since the outset of this chapter, I have emphasized the subjective nature of problem structure. Changing the state of an individual's knowledge concerning the task environment can alter that person's perception of the task (Cronbach & Snow, 1977, pp. 129-130). Hence, the nature of the information provided to the user during training also determines the degree to which the system-use problem is ill-structured. The discussion in this section begins with a summary of selected findings in instructional psychology, followed by a consideration of user training in the MIS/DSS literature.
2.4.1 INSTRUCTIONAL PSYCHOLOGY

Many of the studies in the instructional psychology literature are conducted with children and poor learners, including educable retarded populations. Obviously, the reader has a right to be concerned regarding the generalizability of results to adult users of computer systems. However, the more limited findings based on the behavior of adult subjects (college and above) suggest that "the overall patterns do generalize quite well" (Brown et al., 1983, p. 136). Developmental patterns are echoed in the development of expertise, and areas of developmental differences are also useful in understanding comparative differences in individuals of the same age. For example, field dependent individuals have been described as "developmentally less advanced" (Davis & Frank, 1979). Performance is not always "a function of age per se but of efficient strategy use" (Brown et al., 1983, p. 91) and domain knowledge (Glaser, 1984).

Findings in the instructional psychology literature should be evaluated, however, in terms of differences in objectives and constraints between educational and corporate settings. The issues and tradeoffs are common to both, but costs and benefits
are likely to be estimated differently. In educational settings, instruction is the primary purpose of the organization. Individuals must be prepared in the broadest possible sense for an unknown set of future problem-solving situations. In business settings, time devoted to instruction is weighed against time that could be spent directly meeting the primary purpose of the organization. The set of problem-solving situations the individual will face in the immediate future is more constrained.

One basic tension that might be addressed differently in educational and corporate settings is the use of training resources to develop general or specific skills/knowledge (Brown et al., 1983; Glaser, 1984). The goal of educational instruction is usually "broader than improved performance on a single task" (Brown, 1978, p. 136). In a business setting, improved performance on a recurring, individual task might be a perfectly acceptable objective. This is particularly true since training of more general skills is likely to involve long-term efforts (Pintrich, Cross, Kozma & McKeachie, 1986, p. 641; Frederiksen, 1984, p. 382). Certain options that are available in educational settings, such as avoiding problems beyond students’
current ability, are simply inappropriate in business settings (Case, 1978, p. 221).

One of the most widely-explored phenomena in education is the aptitude-treatment interaction (Cronbach & Snow, 1977). Instructional treatments may have different effects on students depending on their aptitudes, where aptitudes are defined as any characteristic of an individual which predicts learning success. When regression is used to relate performance measures to aptitude levels, aptitude-treatment-interactions (ATI) are expressed as differences in the regression slopes for the treatments. When the regression lines do not cross, an ordinal interaction exists, suggesting that one treatment is more effective than the other regardless of aptitude. Disordinal interactions exist when the regression lines cross, suggesting that aptitude level determines which treatment will be more effective. Taking the differential costs of treatments into account can change the ordinality in either direction. Ordinal and disordinal interactions are graphically depicted in Figure 1.

Intelligence is one of the most important aptitudes to consider in instructional research (Cronbach & Snow, 1977; Pintrich et al., 1986).
Current views of intelligent action stress the importance of cognitive strategies, executive processes and/or metacognitive processes (Brown et al., 1983; Dillon, 1986; Glaser, 1978; Linn, 1986; Pintrich et al., 1986; Snow & Lohman, 1984; Sternberg, 1985). Snow & Lohman (1984) identify fluid intelligence as a primary determinant of aptitude for learning and performance on novel tasks. Fluid intelligence consists of "higher order strategic processes...that make the dynamic cognitive system adapt or learn within a task" (Snow & Lohman, p. 351).

Section 2.3.1 indicated the close conceptual and statistical relationship between tests of CRA and tests of fluid intelligence. Thus, documented ATI involving intelligence are highly relevant to this research. One of the strengths of an ability, as opposed to style, perspective is that these findings can be incorporated rather than ignored.

The characteristic of instruction which most consistently enters into an ATI with intelligence is the structure of the learning situation. Just as the structure of a problem is determined by the extent to which the problem solver must contribute to its solution, an instructional situation can be described as ill-structured when it requires the learner to
engage in "active or complex information processing" (Snow & Lohman, 1984). Structured learning situations are highly organized and complete, demanding little independent inference, organization or reorganization on the part of the learner. Assessments of the results of experimental studies are best made on the basis of a detailed understanding of the nature of the treatments, because the variety of treatment labels used in ATI studies can be misleading (Snow & Lohman, 1984, p. 356).

Advocates of a discovery learning approach suggest that the learning situation should be relatively ill-structured. As the name implies, students are expected to discover general principles with minimal teacher guidance. Discoveries are encouraged through carefully engineered sequences of experiences. Discovery learning is expected to result in a more complete understanding of principles, so that transfer of training is more likely to occur (Greeno, 1978, p. 63; Frederiksen, 1984, p. 397). One danger of discovery learning, however, is that the desired discoveries may not be made at all. Another danger is that the student will make incorrect "discoveries" through an error in reasoning. A practical drawback of discovery approaches in comparison to methods providing more
direct guidance is the comparatively slow pace at which students become able to solve specific problems (Frederiksen, 1984, p. 397).

The consensus on less structured learning situations, including discovery learning environments, is that they favor high ability individuals (Bayman & Mayer, 1988; Brown et al., 1983; Cronbach & Snow, 1977; Gagne & Dick, 1983; Linn, 1986; Sternberg, 1986). Regression slopes become steeper when treatments require elaboration on the part of the learner. The slopes become flatter when "more of the intellectual work is done for the learner" (Cronbach & Snow, 1977, p. 503). This result is consistent with the view that low analytics should benefit more from a structured learning environment (Witkin et al., 1977). However, Halpin & Peterson (1986, p. 972) report that the small number of CRA studies manipulating the instructional environment have yielded inconclusive results.

Intelligence also interacts with treatments which provide direct training in cognitive strategies for learning tasks or ability tests (Snow & Lohman, 1984). In accordance with an increased focus on higher order strategic processes, numerous studies have investigated the possibility of directly training specific strategies in order to improve the performance of
individuals who do not use them spontaneously. Research findings clearly indicate that poor performers can be taught to use strategies similar to those of good performers, with impressive increases in performance (Brown et al., 1983; Case, 1978; Malloy, Mitchell & Gordon, 1987). However, strategy instruction can actually have a negative impact on the performance of high ability subjects when it interferes with existing strategies or is less efficient than existing strategies (Cronbach & Snow, 1977; Snow & Lohman, 1984; Dillon, 1986).

While it is clear that even adults can benefit from explicit instruction in appropriate strategies (Brown et al., 1983; Frederiksen, 1984), there are definite limitations to the approach. First, a detailed understanding of an appropriate strategy must be acquired through task analysis (Brown et al., 1983; Case, 1978; Frederiksen, 1984; Gagne & Dick, 1983; Greeno, 1978). Second, provision of a strategy does not guarantee maintenance or generalization. Maintenance and generalization refer to the application of a strategy over time and across variants of a task, respectively.

It appears somewhat easier to achieve maintenance than generalization of engineered strategies (Brown,
1978). One method for enhancing both maintenance and generalization is to inform subjects of the significance of the strategy in terms of performance (Brown et al., 1983). Sufficient opportunities for practice are also important, since strategies are maintained better when they were executed well at the time of training (Brown et al., 1983; Wexley, 1984).

The transfer of strategies to new tasks is a thorny problem for adults as well as children (Brown et al., 1983). Generalization, like other problem-solving situations, requires flexibility and metacognitive skill (Glaser, 1984; Snow & Lohman, 1984). Greeno (1978) discusses two types of understanding which would appear important to generalization. First, subjects should understand how actions relate to goals and subgoals, a functional or intentional level of understanding. Second, subjects should understand why their actions work in terms of the justifications for each step, an explanatory mode of understanding. Informed instruction could provide this type of information to foster transfer of training (Pintrich et al., 1986). Transfer can also be addressed explicitly as part of cognitive strategy training (Brown et al., 1983).
In summary, the findings in the area of instructional psychology are relevant to corporate settings even though the majority of studies do not use adult subjects. The essential issues are the same whether the instruction takes place in a business or a school, but cost-benefit analysis may lead to different decisions regarding appropriate training objectives. The effects of instructional treatments can be understood in terms of ATI. Low ability subjects are expected to perform better in structured learning situations and when provided with explicit training in appropriate strategies. High ability subjects often perform better in less structured learning situations and without explicit instruction in cognitive strategies. When cognitive strategies are taught explicitly, maintenance and generalization are unlikely to occur without informed instruction and opportunities for practice.

2.4.2 END-USER TRAINING

While the importance of end-user training is widely acknowledged, so is the scarcity of empirical research in the area (Sprague & Carlson, 1982; Bostrom, Olfman & Sein, 1988). Because of the dearth of training studies concerning specific DSS (as opposed to
DSS generators), I have extended the discussion in this section to include research pertaining to a variety of target systems. The need for adequate user task representations (GOMS) provides common ground for all of the studies, although differences in target systems imply differences in the nature and complexity of those representations. The first part of this section is devoted to summarizing existing research, after which the findings are used to support a series of conclusions regarding end-user training.

One training study (Mykytyn, 1988a&b) manipulated target systems rather than training approaches. The target systems differed in terms of the structure of the decision they were designed to support. One DSS supported a short-term cash forecasting decision (structured task), and the other supported a long-term financing decision (semi-structured task). The tasks were appropriate since the subjects were all practicing financial specialists. Training concerning the DSS generator (FCS-EPS) used to create the specific DSS (SDSS) was provided to all users in the same format: classroom instruction (general training) followed by hands-on use of the software's analysis features (task-related training). The SDSS used during training was unrelated to either of the experimental SDSS. The
study explored the relationship between perceptions of training and perceptions and use of the SDSS. Perceptions of training had already been shown to be related to perceived DSS utilization (Fuerst & Cheney, 1982).

One finding was that the pattern of significant relationships differed by task structure. For the structured task, perceptions of both general and task-related training were correlated with perceptions of the contribution of the SDSS. For the unstructured task, only perceptions of the sufficiency of task-related training correlated with perceptions of the contribution of the SDSS.

The other interesting finding was that the SDSS designed to support the structured task was perceived as making a greater contribution to decision effectiveness. The authors explained the result as follows (Mykytyn, 1988a, p. 15):

Possibly less defined problems require a decision maker to exert considerable effort to define the problem and determine appropriate analysis. This may conflict with the use of the...specific DSS...affecting the decision maker’s perception of their value. This result would support the concept that it is important to ensure that DSS and related components provide support in generating solutions and in understanding the problem.
There were no corresponding differences by task structure in users' perceptions of the sufficiency of either form of training.

The simplest contrast between training approaches is the presence or absence of training. An exploratory study by Hughes (1987) investigated this contrast in the context of a financial planning language (or DSS generator). Pretest decision performance (time and number of alternatives considered) was compared to posttest performance in producing a revenue/expenditure estimate. The pretest task was completed manually by both groups, and the posttest task was completed with the DSS generator only for the group which received training. The training was presented in a seminar followed by a "hands-on" workshop. Unfortunately, neither summary data nor a complete analysis were reported. However, the author apparently found no significant performance effects of training or DSS generator use, and virtually no impact of demographic factors (age, sex, education, general and specific programming experience) on the dependent variables.

In a very similar study, Green & Hughes (1986) contrasted the pretest-posttest difference scores for groups receiving four levels of training: no training (and no use of the financial planning language),
seminar only, seminar plus workshop, seminar plus "hands-on" workshop. In this study, the trainee characteristic of interest was an analytic or heuristic cognitive style (as measured by the MBTI). As in Hughes (1987), no hypotheses or even informal predictions were presented concerning the effects of training or cognitive style, nor was a full analysis reported. The result was a somewhat confusing collection of findings which suggested that the extent of training did affect some decision attributes, and that cognitive style affected subjects' response to training. No conjectures as to why the results occurred were provided.

Goslar, Green & Hughes (1986) manipulated DSS training, data level and DSS availability. In this case, however, the SDSS was used through an intermediary so that even subjects without the training session could solve the problem using the SDSS. A number of attributes of subjects' decision-making processes were measured: number of alternatives considered, time, confidence in decision, amount of data considered, perceived method used to make a decision, and performance.

Only the number of alternatives considered was related to any of the independent variables (through
interaction effects involving training). The authors concluded (Goslar et al., p. 88):

Providing such subjects with a decision tool designed for a rational (logical) decision-making process may have been premature. The training provided was nontechnical and presented the features and potential applications of a DSS; it provided a rational approach to problem solving using the DSS features available. For individuals not accustomed to approaching and solving ill-structured problems systematically, training of this type may not have provided the information necessary.

They go on to suggest that "in the future, DSS training may need to include decision-making and logic training in addition to technical instruction if users are to take full advantage of DSS capabilities" (Goslar et al., 1986, p. 88). It should be noted that the subjects in this experiment were marketing employees solving a marketing case study, so the results cannot easily be attributed to a lack of relevant knowledge concerning the problem domain. This study calls attention to the difficulty of formulating appropriate GMS (goals, methods and selection rules), since the details of operator use were the responsibility of intermediaries.

Although not as clearly a training study, the work of Cats-Baril & Huber (1987) is highly relevant to the
question of decision-making training. The independent variable of most interest to end-user training was the presence or absence of a decision-aid heuristic (attributes of the DSS were also manipulated). Although the heuristic was provided only through written instructions as part of the DSS, the same message could just as easily have been conveyed in a more "traditional" training session. The outcomes can be examined in terms of performance variables and affective variables.

Individual differences were not found to influence any performance (or affective) outcomes. The experiment did show that subjects provided with the heuristic performed better than subjects without the heuristic in terms of the quality of career plans they developed and their productivity of ideas. "These are important findings given the emphasis of the information systems field on providing data rather than decision rules to support users" (Cats-Baril & Huber, 1987).

While subjects provided with the heuristic performed better, they were also less satisfied with their decision processes, were less confident in their decisions, and showed a negative shift in attitudes towards the importance of career planning. These negative responses were interpreted, however, as
"positive signs of introspection and of effective searching for information" in an ill-structured task (Cats-Baril & Huber, 1987, p. 368). The same inverse relationship between training procedure effectiveness and user perceptions was reported by Smith & Kight (1959). Among other measures, the authors suggested that the affective outcomes might have been improved by training users in the heuristics before they interacted with the system and/or by making the support system more flexible, in appearance or in fact (Cats-Baril & Huber, 1987, p. 370).

Rather than contrasting the presence or absence of training, an early work conducted by Grace (1977) compared two types of training. Because of the high cost of instructor time, especially for one-on-one tutorials, it was important to establish that other training approaches could be equally effective. Grace developed a computerized interactive training manual (ITM) to train users of GADS.(geo-data analysis and display system). Outcomes (instructor time, test performance, student time, student motivation) for subjects learning through use of the ITM with instructor assistance were compared to outcomes for subjects learning through one-on-one tutorials.
On the average, subject performance and motivation were equivalent for either training approach, while instructor time was much lower for the ITM group. Because instructors in the tutorial group were "allowed to use any teaching techniques or learning sequences they desired, including the sequence in the ITM" (Grace, 1977, p. 35), it is impossible to know in any detail how the training approaches compared. If one assumes that the approaches were not drastically different, then this study confirms a conclusion reached by instructional psychologists: "one medium is about as good as another" (Pintrich et al., 1986). In this case, the particular training approach embedded in the ITM was as effective an approach as those adopted by the human tutors, at least with respect to the comprehension test administered. A fundamental, unanswered question remains: what makes a training approach effective?

Bostrom, Olfman & Sein (1988) suggest that training environments can be viewed from two perspectives: physical and methodological. Physical characteristics are essentially attributes of the messenger rather than the message. For example, Sprague & Carlson suggest six physically different categories of user-education techniques: tutorial,
course/seminar, programmed instruction/computer-assisted instruction, interactive training manual, local expert, and DSS help component (1982, p. 156). A long list of other physical characteristics could be identified (e.g., duration, timing, location). In comparison, training methods "refer to strategies that are employed in instructing the learners" (Bostrom et al., 1988, p. 233).

Bostrom et al. proceed to echo the conclusion presented by Pintrich et al.: "physical environment design is important, but it will depend on how instructional methods are embedded in the design" (Bostrom et al., 1988, p. 233). Their research draws on work in mental models and aptitude-treatment interactions in an effort to determine what methodologies work best and for whom. Their work differs from my own in focusing on different aspects of mental models and different individual differences.

One study contrasted IFPS training which explicitly provided either an analogical model, an abstract model, or no conceptual model at all (Olfman, Sein & Bostrom, 1986). The analogical model treatment described IFPS as a spreadsheet, while the abstract model treatment described it in terms of mathematical relations (Olfman et al., 1986, p. 5). Student
comprehension was measured after training with a test instrument "designed to capture students' knowledge about the IFPS language rather than their ability to solve problems using IFPS" (Olfman et al., 1986, p. 6). Given the nature of the test, performance reflected primarily the operator portion of students' mental models.

The results showed that individual differences in a number of basic abilities explained very little of the variance in performance. I would suggest that this result may have been due to the nature of the subject population and/or the context in which performance was measured. The mean scores for the subject population were quite high, suggesting that there might not have been enough variation in abilities to affect performance. Some abilities might also have had a greater relationship to problem-solving performance than to what was essentially recall performance. As mentioned earlier, the formulation of GMS was not required by the testing situation.

There was almost no difference in comprehension performance (operator knowledge) between treatments. The authors suggested that this may have been due to weak treatments which did not sufficiently stress the conceptual model. Another possibility is that all of
the subjects were experienced learners (graduate students) of fairly high abilities capable of adapting to several modes of instruction. Halpin & Peterson (1986) presented this as a plausible explanation for the lack of an interaction in their own research on instructional treatments.

Sein & Bostrom (1989) conducted another study comparing analogical and abstract conceptual models. The target system for this study was VAX MAIL. The analogical model explained the hierarchical structure of the system in terms of an office filing cabinet. The abstract model simply presented a diagram showing the hierarchical relationship of labelled rectangles representing the objects making up the system. While this study neither predicted nor found main effects of model type, interaction effects were discovered.

The expected aptitude-treatment interactions were based on the relationship of selected user characteristics to the processes of constructing and "running" a mental model. Analogical and abstract models were viewed as two ends of a continuum which are distinguished by the ease of mapping the information contained in the model to the target system. The relationship of objects in an analogical model to objects in the target system is more easily
appreciated, requiring less effortful processing on the part of the learner.

Visual ability was seen as a relevant measure of an individual's skill in both performing the necessary mapping activities and manipulating the resulting representation. Hence, high visual ability subjects were expected to be able to utilize either type of conceptual model. Low visual ability subjects, on the other hand, would be unable to use the more difficult abstract conceptual model.

The study results confirmed the authors' predictions. High visual ability subjects performed better than low visual ability subjects overall. There was an interaction of model type and visual ability \((p=.072)\) for the more complex experimental tasks (demanding the formulation of GMS), but not for simpler tasks or a comprehension test. Performance at both ability levels was equivalent for the analogical model treatment, but high visual ability subjects outperformed low visual ability subjects with the abstract model.

The abstract versus concrete dimension of learning mode as measured by the Learning Style Inventory (Kolb, 1976) was also considered with respect to aptitude-treatment interactions. Kolb's learning theory states
that abstract learners prefer an analytical approach to learning, while concrete learners prefer using prior referent experiences (Kolb, 1971). Hence, abstract models should appear more attractive to abstract learners, and analogical models should be preferred by concrete learners. The authors suggest that the preferred model type suggested by theory will also lead to better performance. The line of reasoning behind the predicted interaction, however, depends on experience and ability differences rather than mere preference differences. The authors have acknowledged elsewhere that there is an imperfect correspondence between stated preferences as measured by the KLSI and abilities (Bostrom, Olfman & Sein, 1990, p. 109).

The authors’ predictions were also confirmed with respect to learning mode. Abstract learners performed better than concrete learners for all performance measures. Again, an interaction was present only for the complex experimental tasks. Concrete learners outperformed abstract learners in the analogical model treatment, and abstract learners outperformed concrete learners in the abstract model treatment. Individuals must have lacked the desire and/or skills necessary to utilize information presented in the non-preferred mode.
The final study considered in this section featured another type of methodological comparison: construct-based training versus applications-based training (Olfman, 1987; Olfman & Bostrom, 1988). These terms refer to essentially the same methods as Ginzberg's "operations training" and "task context training" (1978, p. 49). The construct-based training emphasized features and commands of the software and used generic problem-solving exercises. The applications-based training emphasized problem solving, including the solution of problems which trainees encountered on the job. The target software in this study was Lotus 1-2-3. Since both treatments included problem-solving exercises, all subjects had an opportunity to acquire mental models beyond the operator level. However, the GMS acquired during each treatment were formulated on the basis of different experiences (different exercises).

The argument for the superiority of applications-based training was based on the concept of personal relevance. The argument was that more personally relevant information would be better assimilated into existing knowledge structures and produce superior motivation. Overall, however, the differences between training methods as measured by a comprehension test
and a budgeting task were not significant. Application-based training did lead to higher self-reported hours of use eight weeks subsequent to training.

2.4.3 CONCLUSIONS REGARDING END-USER TRAINING

The studies reviewed in the previous section are summarized in Table 3. The research findings presented above can be used to support the following conclusions:

Conclusion 1. Individual differences are important in the study of end-user training.

One study showed that learning mode was related to performance outcomes (Sein & Bostrom, 1989). Another showed that cognitive style interacted with training (Green & Hughes, 1986). Besides demonstrating that individual differences are important, the studies confirm the usefulness of an aptitude-treatment interaction perspective. Interestingly, Sein and Bostrom's interpretation of visual ability (1989, p. 206) is very similar to the interpretation of cognitive restructuring ability presented earlier in this chapter.

There were also studies in which individual differences did not appear to influence the response to training (Hughes, 1987; Olfman, Sein & Bostrom, 1986;
Cats-Baril & Huber, 1987). Explanations were not always provided as to the relevance of the individual difference variables selected or the mechanisms through which they would interact with training. Hence, it is difficult to interpret this lack of results. Conclusion number three below may have some bearing on why performance did not necessarily reflect individual differences.

Conclusion 2. Users could benefit from assistance with problem structuring, particularly in less-structured environments.

This was demonstrated in Cats-Baril & Huber’s study (1987) where availability of a heuristic was associated with improved performance. Based on the results of their own studies, the same conclusion was reached by Mykytyn (1988a,b) and Goslar, Green & Hughes (1986). Both Ginzberg (1978) and Hellman (1989) have suggested that knowledge of tasks (GMS) should not be overlooked during end-user training.

Conclusion 3. The training methodologies studied have had surprisingly little differential effects on performance.

I believe that the main reason for the lack of strong effects was the lack of attention to process. Except for Cats-Baril & Huber (1987) the treatments considered did not explicitly address the area in which
strong performance improvements have been documented: specific cognitive strategies. For the most part, they represented discovery learning environments in which the users had to inductively derive principles from examples. Knowledge of operators was provided, but information concerning appropriate GMS was limited. Strategies determine performance, but the treatments did not differ in the extent of strategic support provided (with the exception noted above).

The GOMS framework could contribute substantially towards interpreting past results and designing future research. The objective of training is to enhance users' mental models, but we should clarify what aspect of the mental model is being addressed by a training method in terms of the GOMS framework. For example, the comparison of abstract and analogical conceptual models should take into account how the models assist in developing GOMS. In Olfman et al. (1986), the impact of both conceptual models, if any, would seem to be in assisting the subject to organize knowledge of operators. In Cats-Baril & Huber (1987), one treatment provided only operators whereas the other treatment provided GOMS. Methods for conveying the same sort of assistance will probably result in smaller performance
differences than methods which convey different types of assistance.

The design of testing should, of course, reflect the nature of expected treatment differences. The observed pattern of ATI can depend on the criterion against which training is evaluated (Brown et al., 1983, p. 138). Tests can be understood in terms of the demands they place on GOMS.

The study by Sein & Bostrom (1989) illustrates this point. Interaction effects were observed only for the complex (far-transfer) tasks, not for the simple (near-transfer) tasks or a comprehension test. More complex problem-solving tasks make demands on the higher levels of subjects' mental models (GMS) and/or on their problem-solving skills. Differences at these levels will not be revealed by tests of operator knowledge or simple tasks requiring trivial GMS. Of course, this point is obvious when user models, the information conveyed during training, and task demands are understood in terms of the GOMS viewpoint.

The study by Olfman & Bostrom (1986) also illustrates the need to exercise care in establishing performance criteria. As discussed above, the experimental treatments differed in the set of exercises used to develop GMS knowledge. Applications-
Based training enhanced the particular GMS associated with on-the-job problems. Construct-based training developed GMS in the context of generic problems. While both treatments addressed the range of GOMS, there is no reason to believe that the set of GMS developed in one context is equivalent to the set of GMS acquired in another problem context. Since subject performance was assessed with respect to yet a different problem-solving context, neither set of GOMS may have been adequate (or either set of GOMS may have been adequate).

The real advantage of applications-based or task-context training is more likely in ensuring that the set of GOMS is appropriate for on-the-job problems, but on-the-job performance testing could not be attempted given the design of the study. Using an on-the-job criterion, self-reported hours of use were higher for applications-based training. Perhaps the conclusion regarding performance would have been different if an on-the-job performance criterion had been used as well.

Mykytyn (1988a&b) also measured performance in a context which differed from the training context (completely unrelated SDSS were involved). Whenever the training context differs from the testing context (with respect to the system or task), transfer of
training must occur in order for the training to produce positive effects. As discussed in Section 2.4.1, training often fails to generalize even for adult learners. Training methods which are indistinguishable in terms of the success of transfer may have differential effects on performance as measured in a problem-solving context similar to the training context (and vice versa). Training situations which more closely resemble testing situations are likely to produce better results because the need for transfer is reduced.
### List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATI</td>
<td>Aptitude-Treatment Interaction</td>
</tr>
<tr>
<td>CRA</td>
<td>Cognitive Restructuring Ability (measured by GEFT)</td>
</tr>
<tr>
<td>DSS</td>
<td>Decision Support System</td>
</tr>
<tr>
<td>GEFT</td>
<td>Group Embedded Figures Test</td>
</tr>
<tr>
<td>GMS</td>
<td>Goals, Methods and Selection Rules (see GOMS below)</td>
</tr>
<tr>
<td>GOMS</td>
<td>Goals, Operators, Methods and Selection Rules (the structure of the user's strategy)</td>
</tr>
<tr>
<td>HO</td>
<td>High analytic subject in Operations training treatment</td>
</tr>
<tr>
<td>HS</td>
<td>High analytic subject in Strategy training treatment</td>
</tr>
<tr>
<td>LO</td>
<td>Low analytic subject in Operations training treatment</td>
</tr>
<tr>
<td>LS</td>
<td>Low analytic subject in Strategy training treatment</td>
</tr>
<tr>
<td>MIS</td>
<td>Management Information System</td>
</tr>
<tr>
<td>SDSS</td>
<td>Specific Decision Support System (as opposed to a DSS generator)</td>
</tr>
</tbody>
</table>
### Table 2 (continued on next page)

**Summary of CRA Studies in an MIS/DSS Context**

<table>
<thead>
<tr>
<th>Reference</th>
<th>System Design Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lusk (1979)</td>
<td>Compared two tabular and three graphical report formats.</td>
</tr>
<tr>
<td>Lusk &amp; Kersnick (1979)</td>
<td></td>
</tr>
<tr>
<td><strong>Findings:</strong> Higher performance with less complex (tabular) report formats; Higher performance by high analytics; No differential performance peaking.</td>
<td></td>
</tr>
<tr>
<td><strong>Findings:</strong> More high analytics completed the task successfully; High analytics used more graphical reports (the report judged more relevant by the authors).</td>
<td></td>
</tr>
<tr>
<td>Benbasat &amp; Dexter (1985)</td>
<td>Compared multi-color and mono-color reports in either tabular or graphical format.</td>
</tr>
<tr>
<td><strong>Findings:</strong> High and low analytics performed equally with multi-color reports; High analytics performed better than low analytics with mono-color reports; Low analytics rated tabular report as more understandable (no difference for high analytics).</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** The summary provided in the table has been limited to those aspects of the studies relevant to the discussion of CRA in Section 2.3.2.
Table 2 (continued from previous page)

Summary of CRA Studies in an MIS/DSS Context

<table>
<thead>
<tr>
<th>Reference</th>
<th>System Design Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benbasat &amp; Dexter</td>
<td>Compared database inquiry capability with structured/aggregate report.</td>
</tr>
<tr>
<td>(1979)</td>
<td></td>
</tr>
<tr>
<td><strong>Findings</strong>: High and low analytics performed equally with database inquiry capability; High analytics performed better than low analytics with the structured/aggregate report; Some differences in report request behavior; High analytics reported a more stable, fixed strategy approach.</td>
<td></td>
</tr>
<tr>
<td>Benbasat &amp; Dexter</td>
<td>Compared presence/absence of a decision aid (used the structured/aggregate report from Benbasat &amp; Dexter [1979]).</td>
</tr>
<tr>
<td>(1982)</td>
<td></td>
</tr>
<tr>
<td><strong>Findings</strong>: The decision aid resulted in better performance for both high and low analytics, but the difference between treatment groups was much greater for low analytics.</td>
<td></td>
</tr>
<tr>
<td>Pracht &amp; Courtney</td>
<td>Compared presence/absence of a graphical decision aid.</td>
</tr>
<tr>
<td>(1988)</td>
<td></td>
</tr>
<tr>
<td><strong>Findings</strong>: For high analytics, decision aid group performed better. For low analytics, no difference in performance associated with presence of decision aid.</td>
<td></td>
</tr>
</tbody>
</table>

Note. The summary provided in the table has been limited to those aspects of the studies relevant to the discussion of CRA in Section 2.3.2.
Figure 1. Examples of ordinal and disordinal interactions.
Table 3 (continued on next page)

**Summary of MIS/DSS Training Studies**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Treatment Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mykytyn (1988a&amp;b)</td>
<td>SDSS designed to support structured or semi-structured task (training for all subjects provided during seminar &amp; &quot;hands-on&quot; workshop).</td>
</tr>
<tr>
<td><strong>Findings:</strong></td>
<td>Pattern of correlations depended on structure of task; Perceptions of DSS contribution to effectiveness lower for semi-structured task.</td>
</tr>
<tr>
<td><strong>Findings:</strong></td>
<td>Training had no effect on performance variables; Demographic variables (age, sex, experience) had little relationship to performance.</td>
</tr>
<tr>
<td>Green &amp; Hughes (1986)</td>
<td>Compared four levels of training: no training, seminar, seminar &amp; workshop, seminar &amp; &quot;hands-on&quot; workshop.</td>
</tr>
<tr>
<td><strong>Findings:</strong></td>
<td>Extent of training affected some decision attributes; cognitive style affected subjects' response to training.</td>
</tr>
</tbody>
</table>

**Note.** The summary provided in the table has been limited to those aspects of the studies relevant to the discussion of training in Section 2.4.2.
### Table 3 (continued on next page)

#### Summary of MIS/DSS Training Studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Treatment Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goslar, Green &amp; Hughes (1986)</td>
<td>Contrasted presence/absence of training (training provided during seminar); All subjects used SDSS through intermediaries.</td>
</tr>
<tr>
<td><strong>Findings:</strong></td>
<td>No significant differences in decision-making confidence, decision-making processes, or performance levels.</td>
</tr>
<tr>
<td><strong>Findings:</strong></td>
<td>Performance better for groups provided with heuristic, but use of heuristics associated with negative affective outcomes.</td>
</tr>
<tr>
<td>Grace (1977)</td>
<td>Contrasted one-on-one tutorial with ITM (Interactive Training Manual)</td>
</tr>
<tr>
<td><strong>Findings:</strong></td>
<td>Subject performance and motivation were equivalent for either training approach; Instructor time was much lower for ITM group.</td>
</tr>
</tbody>
</table>

**Note.** The summary provided in the table has been limited to those aspects of the studies relevant to the discussion of training in Section 2.4.2.
Table 3 (continued from previous page)

**Summary of MIS/DSS Training Studies**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Treatment Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olfman, Sein &amp; Bostrom (1986)</td>
<td>Compared training based on an analogical model, abstract model, or no conceptual model.</td>
</tr>
<tr>
<td><strong>Findings:</strong> Basic abilities explained a significant but small amount of the variance in performance; No significant performance differences between training treatment groups.</td>
<td></td>
</tr>
<tr>
<td>Sein &amp; Bostrom (1989)</td>
<td>Contrasted training based on an analogical model or an abstract model.</td>
</tr>
<tr>
<td><strong>Findings:</strong> Interaction effects of training and individual differences (visual ability, learning style) for far-transfer tasks but not near-transfer tasks; No main effect of training type.</td>
<td></td>
</tr>
<tr>
<td>Olfman (1987)</td>
<td>Contrasted applications-based and construct-based training (applications-based training used job-specific training exercises and construct-based training used generic training exercises.).</td>
</tr>
<tr>
<td>Olfman &amp; Bostrom (1988)</td>
<td><strong>Findings:</strong> Main and interaction effects of learning style; No main effect of training method; Self-reported hours of system use following training higher for applications-based treatment.</td>
</tr>
</tbody>
</table>

**Note.** The summary provided in the table has been limited to those aspects of the studies relevant to the discussion of training in Section 2.4.2.
CHAPTER 3. RESEARCH FRAMEWORK AND DESIGN OVERVIEW

This chapter presents the research framework and research questions guiding this study. Predictions are made based on the research findings and conclusions presented in the previous chapter. A brief summary of the experimental design is provided here, and details of the research design and hypotheses are presented in Chapter 5.

3.1 RESEARCH FRAMEWORK

The research framework guiding this study is presented in Figure 2. Each component of the framework has been numbered, and each link can be referenced by the components it connects (e.g., 1-2 indicates the link from component 1 to component 2). The framework is organized around the distinction between the strategy formulation process (2) and the strategy execution process (4).

During the strategy formulation process (2), the subject selectively attends to information available from a variety of sources. The external environment includes information available via training (6,7), the system interface (8), and task characteristics (9). The subject’s internal environment includes domain
knowledge (10). Once information has been selected, it is organized to facilitate storage, later retrieval and use. The information is reorganized as necessary to create a set of GOMS (3). Cognitive restructuring ability (1) contributes to the strategy formulation process (1-2) because the process requires the skills identified with CRA: selecting, organizing and reorganizing information (as discussed in Section 2.3.1).

During the strategy execution process (4), the subject retrieves information concerning GOMS (3), performs the necessary operations, and processes feedback. Outcome feedback may lead to changes in the execution process (5-4). Feedback (5-4) may also prompt renewed strategy formulation (4-2) by providing information gained from the "iterative taking of action" (Connolly, 1982). Limited evidence exists to suggest that CRA affects the strategy execution process (1-4), as discussed in Section 2.3.1.

3.2 DESIGN OVERVIEW

In this study, the structure of the learning situation is manipulated by controlling the amount of structuring support provided during training (6,7). A training manipulation, as opposed to an interface
design manipulation, is the focus of study. The help function available through the system interface (8) is manipulated to match the training treatment, but all other aspects of the interface are held constant. CRA (1) is measured, task characteristics (9) are held constant, and all subjects receive the same lecture concerning the domain, although domain knowledge (10) is not necessarily constant. As discussed in the next chapter, the experimental DSS interface (8) contributes little structure (is highly flexible), and the user is faced with a semi-structured task (9). The outcome variables (5) studied include performance, decision time, decision confidence and perceptions of the system. Attributes of user strategies (3) are also investigated.

A GOMS framework can be used to describe the two types of training considered: strategy training (6) and operations training (7). Strategy training (6) provides integrated information concerning GOMS to the subject. Operations training (7) provides information concerning operators only. The information conveyed during training is subject to the same processes of selection, organization and reorganization as any other form of information available to the subject. Hence, CRA affects the subject’s response to training (1-2)
just as it affects the subject’s response to any other form of information.

The choice of training treatments was made based on a number of considerations. Research in instructional psychology has demonstrated that strategy instruction can produce substantial gains in performance, particularly for lower ability subjects (see Section 2.4.1). While MIS/DSS research has shown that users could benefit from assistance with problem structuring, the extent of structuring support provided during training has rarely been manipulated (see Section 2.4.3).

The selection of CRA as the individual difference measure is likewise supported by research findings. The role of CRA in structuring problems makes it relevant to the strategy formulation process (see Section 2.3.1). Its close relationship to fluid intelligence and the observed interactions between intelligence and extent of strategic support also suggest the relevance of CRA to the training treatments chosen for this study (see Section 2.4.1). CRA has already been linked to performance differences in an MIS/DSS context (see Section 2.3.2).
3.3 RESEARCH QUESTIONS AND PREDICTIONS

As discussed above, a number of outcome variables are measured (performance, decision time, decision confidence and perceptions of the system). However, because this work is concerned with decision effectiveness, performance is the outcome variable of primary interest. Additionally, adequate research bases for predictions exist only with respect to performance. For these reasons, only research questions related to performance are stated below. Chapter 5 provides a discussion of other outcome variables.

There are a number of research questions addressed by this study:

1) **Does the training treatment (6,7) alter the relationship between CRA (1) and performance (5)?**

High analytics are expected to outperform low analytics in the operations training treatment (see Section 2.3.3). The experimental task is a novel, semi-structured, analytical, problem-solving task which should place high demands on CRA. Operations training provides the minimal information necessary to formulate a decision strategy incorporating the computerized system.
CRA is not expected to be associated with performance for strategy-trained subjects. Instruction in cognitive strategies has improved the performance of low ability populations (see Section 2.4.1). ATI research, including some research directly concerning CRA (Witkin et al, 1977; Halpin & Peterson, 1986, p. 972), suggests that more structured learning situations should produce higher performance for low analytics by reducing the demand for structuring skill. The types of interactions observed in CRA research in an MIS/DSS context (Section 2.3.2) also suggest that low analytics can perform as well as high analytics under appropriate conditions.

Using strategy training to improve the performance of low analytics is an example of an approach called compensation (Cronbach & Snow, 1977). Because there is very limited evidence (Witkin & Goodenough, 1981) that we can improve CRA in general (an approach called remediation), we must find ways to compensate for weak cognitive restructuring skills. We can, in certain circumstances, reduce the demands we place on CRA with respect to a specific task. We can attempt to simplify the problem posed by a system to its user. The method explored in this research is to augment the user’s strategy formulation process by providing strategy
instruction with respect to a specific task. Difficulties in the process of mapping via usage, particularly for a novel task and flexible DSS, are anticipated and offset through training.

Whether low analytics do perform as well as high analytics will depend on whether low analytics: 1) learn the strategy equally well, 2) execute the strategy equally well, and 3) are equally willing to adopt the strategy. The impact of strategy execution is explored below in Research Question 1.2. Regarding the third condition, evidence indicates that the behavior of low analytics is quite malleable by "as simple a means as providing directions" (Witkin et al., 1977, p. 26). Also, all of the subjects in the pretest attempted to use the suggested strategy.

The impact of differences in learning during strategy training are anticipated to be minor for three reasons. First, all of the subjects should be fairly expert learners, given that they are adult graduate students. Second, difficulties in the process of mapping via training are considered in the design of instruction, so that a highly structured environment compensates for low structuring abilities. The information provided during training is already organized in terms of GOMS, and requires no
reorganization in order to use. Finally, the system help function can be used to "relearn" the strategy as necessary.

Because the origin of the relationship between CRA and performance is unclear (see Section 2.3.1), two secondary research questions are explored:

1.1) What is the relationship between CRA and the strategy formulation process (1-2) as reflected in the nature of user strategies (3)?

The strategies of high analytics are expected to be more analytical than the strategies of low analytics (an "analytical problem-solving approach" is operationalized in Chapter 5). Some strategy differences have been documented (see Section 2.3.1). Differences in the nature of user strategies are also expected to produce differences in the patterns of information use. Because of their greater skill in separating relevant from irrelevant information, high analytics should use more of the information "relevant" to an analytical problem-solving approach (Relevance is also operationalized in Chapter 5.).

1.2) What is the relationship between CRA and the strategy execution process (1-4)?

Low analytics are expected to experience greater difficulty during the strategy execution process
("Execution difficulty" is operationalized in Chapter 5.), although the nature of the difficulty may depend on the experimental treatment. Low analytics might experience greater difficulty because of the slower automatization of their strategies (see Section 2.3.1), particularly in the strategy training condition. Strategy-trained low analytics might also experience difficulty if the trained strategy is "unnatural" for them, although the relationship of decision model "naturalness" to performance is unreliable (Kottemann & Remus, 1989). Due to weaknesses in strategy formulation, low analytics might find their set of GOMS inadequate in terms of completeness or performance, particularly in the operations training treatment.

The second primary research question concerns the main effect of training content:

2) What is the relationship between training treatment (6,7) and performance (5)?

Strategy-trained subjects are expected to outperform operations-trained subjects. The instructional psychology literature suggests that substantial performance benefits can be obtained by training specific cognitive strategies, even when the
strategies would not have been created/applied spontaneously (see Section 2.4.1).

This prediction is, of course, based on the assumptions that the strategy presented during training: 1) is superior to at least some of the strategies developed by operations-trained subjects, and 2) can be learned and applied by at least some of the subjects. Since the suggested strategy was developed with expert input over an extended period of time, there was reason to believe that it would be superior to the strategies developed by novices in a relatively short period of time. In fact, pretesting showed that both of these assumptions are met with respect to the suggested strategy.

Assuming that the prediction concerning the aptitude-treatment interaction is confirmed, an ordinal interaction is expected. The interaction will be ordinal, or nearly so, if: 1) the trained strategy is as effective and efficient as the majority of spontaneously devised strategies, and 2) the strategy does not interfere with the subject’s existing strategies. As discussed above, there is reason to believe that the suggested strategy is superior. Because the task is novel and the training is provided
prior to any independent strategy formulation attempts, interference with existing strategies is not anticipated.

An ordinal interaction suggests that, on the basis of performance alone, one instructional treatment could be provided to all users regardless of CRA. This would be a practically important result because it would obviate the need for psychological testing prior to training and reduce the cost of developing instructional materials. A disordinal interaction would require a more careful consideration of the appropriate response.

Assuming that a relationship is found between training treatment and performance, the origins of the difference can again be explored in terms of the strategy formulation and strategy execution processes:

2.1) What is the relationship between training treatment and the strategy formulation process (6,7-2) as reflected by the nature of user strategies (3)?

Strategy differences should exist between the treatment groups. The nature of the suggested strategy is known. Differences in the nature of user strategies are also expected to produce differences in the pattern of information use. Strategy-trained subjects should be more likely than operations-trained subjects to use
the items required by the suggested strategy, and less likely than operations-trained subjects to use the other information items available from the system (Exceptions are noted in Chapter 5).

2.2) What is the relationship between training treatment and the strategy execution process (6,7-4)?

This question is purely exploratory, and no predictions are made concerning the nature of the relationship, if any.

In summary, this study focuses on the combined effects of training treatment and CRA on user performance. The expected nature of the aptitude-treatment interaction has been specified. This study also examines the origin of performance differences in terms of the strategy formulation process, the strategies which are created during that process, and the strategy execution process.
Figure 2. Research framework.
CHAPTER 4. TASK, DSS AND STRATEGY DEVELOPMENT

This chapter provides a detailed description of the task environment and decision support system which were common to all subjects. The development and final form of the suggested strategy are also discussed. The material presented in this chapter is useful in appreciating the nature of the experiment and in understanding the data collected, particularly verbal protocol data.

4.1 TASK SELECTION PROCESS

Given the research questions presented in the last chapter, I identified four characteristics of an ideal experimental task:

1) an objective performance score comparable across subjects as a measure of decision making performance,

2) sufficient demands on the strategy formulation process to require cognitive restructuring ability and create a range of performances,

3) sufficient complexity in the nature of a successful problem-solving strategy such that strategy instruction would be nontrivial, and

4) the ability to compare results to other GEFT research within the MIS/DSS literature where performance differences related to CRA have been reported.
Practical constraints dictated by the limited availability of subjects also entered into the selection process.

The obvious starting point in the selection of the experimental task was to evaluate the tasks meeting the fourth criterion, comparability to past GEFT studies. Evaluations of the tasks used in the studies reviewed in the second chapter are presented in Table 4. Each of the tasks is rated with respect to feasibility and the first three criteria listed above. Each evaluation is explained below.

The first task (Benbasat & Dexter, 1979, 1982) was originally evaluated as satisfactory in all respects and was selected as the experimental task. The task environment "was a relatively structured multiperiod inventory/production decision-making game" (Benbasat & Dexter, 1982, p. 2). Subjects consulted an on-line aggregate report in order to set order points, order quantities and production figures for a hypothetical firm. The objective was to maximize "the total profit of the firm over the duration of the game" (Benbasat & Dexter, 1982, p. 2).

When domain experts were consulted regarding an appropriate strategy, however, their responses indicated that the nature of a suggested strategy would
be extremely simple. The strategy would be centered around a demand forecast and application of the economic order quantity formula. One expert suggested that the same decision should then be repeated for several periods without reference to the report provided by the system! The responses of domain experts suggested that strategy instruction would be trivial for this task.

The second task (Benbasat, Dexter & Masulis, 1981) was questionable with respect to the demands it placed on the strategy formulation process, and unsatisfactory with respect to the complexity of a successful strategy. Subjects were required to locate the price/quantity combination which would maximize profit for a single product. System output was provided in three formats. Strategy training would require only informing the subjects of the most appropriate format, assuring that they could interpret the output, and suggesting a search pattern to be followed.

The third task (Benbasat & Dexter, 1985) failed to meet the criterion of nontrivial instruction. Subjects were required to allocate a promotional budget across three territories using a tabular or graphical report. Only one paragraph was required to summarize a reasonable strategy for the budget allocation problem
(Benbasat & Dexter, 1985, p. 1354). Additionally, the performance differences associated with CRA in this study appeared to reflect differences in visual perception (the extent to which color aided disembedding) rather than differences in the strategy formulation process.

Dos Santos (1982) provided subjects with a simulated business environment which could be manipulated to gain an understanding of the relationships between variables. In order to assist subjects in acquiring that understanding, they would have to be instructed in a simple method for systematically testing variable relationships. Understanding was measured in terms of subjects' ability to apply their knowledge to finding problems, prioritizing problems, and finding an alternative course of action. It is impossible to determine how much task performance depended on an understanding of the relationships versus an application of that understanding. The complexity of a suggested strategy was evaluated as marginal rather than unsatisfactory because the strategy might have to address the application of subjects' knowledge in addition to helping them acquire that knowledge.
The fifth task (Lusk, 1979; Lusk & Kersnick, 1979) relied heavily on mathematical structuring activity in the form of calculation. Subjects were required to answer questions using variously transformed versions of a report. Some questions could be answered directly from the report, but 80% of the items required that a transformation be performed on the data. While I agree with the authors that it may be possible to use training to eliminate performance differences in this context (Lusk & Kersnick, 1979, p. 797), the mathematical strategies required are quite simple. The purely mathematical nature of the required strategies would also limit the generalizability of the results.

The final task (Pracht & Courtney, 1988) was infeasible due to time limitations. As in the DosSantos study (1982), the task consisted of using a simulated business environment in order to gain an understanding of the relationship between variables and then applying that knowledge in answering a series of questions. The authors were able to demand a considerable time investment from their subjects outside of the classroom because the inducements were considerable (exemption from a final exam or term paper worth 20% of the course grade). In comparison, the instructors willing to make their students available to
me did not wish the experiment to significantly affect course grades and offered limited access during class time. I would have been unable to secure enough participants willing to make the time investment demanded by the task outside of class time.

In view of the weaknesses summarized in Table 4, a new task environment was developed at the expense of comparability. Because existing task environments were often weak in terms of strategy formulation demands and the complexity of a successful strategy, a task was selected which is unstructured in a number of ways even for domain experts. First, there is no established strategy for solving the problem. Even though the experimental task is an example of the simplest kind of scheduling problem, there is no widely accepted heuristic to meet the objective of minimizing total cost. Second, the search space is quite large (7! permutations for seven jobs). Third, the task is completely novel to the experimental subjects, although they have been exposed to the necessary background information. Finally, there is no definite criteria by which an individual decision maker can determine how well the problem has been solved. All of these characteristics have been associated with unstructured

4.2 TASK DESCRIPTION

The task comes from the domain of job scheduling, a topic typically included in introductory operations management texts. The experimental task is a one-machine sequencing problem in a static and deterministic setting with no preemption. For the experiment, the subject specified an order in which to complete seven jobs waiting for processing. The known fundamental characteristics of each job are: required periods of processing time (duration), period due, in-process inventory holding cost per period, tardiness cost per period, and the nonsymmetric changeover cost to any other job.

The objective is to create a sequence which minimizes the sum of in-process inventory holding charges, tardiness charges and changeover charges. The complexity of the problem derives from the objective of minimizing the sum of three types of charges. Relatively simple heuristics have been developed for the more structured problem of controlling one category of charges at a time, even though these heuristics do not necessarily guarantee minimums (Emmons, 1984).
While the problem is challenging even for a domain expert, the basic knowledge required to function rationally is simple enough that novices can be expected to grasp the essential relationships in a relatively short period of time. The rest of this section briefly summarizes the concepts which are presented to subjects prior to their problem-solving efforts. The presentation to subjects also includes some background information on job scheduling and examples of computations and heuristics as applied to specific sequences.

4.2.1 CONTROLLING IN-PROCESS INVENTORY HOLDING CHARGES

When all other job characteristics are equivalent, jobs which require longer processing times should appear later in the sequence. This is known as the shortest processing time rule. When all other job characteristics are equivalent, jobs with higher in-process inventory holding costs per period should appear earlier in the sequence. Consideration of processing times and inventory holding costs per period can be combined in a single rule known as Smith’s rule.
Using Smith’s rule, jobs are sequenced in increasing order of:

\[
\frac{(\text{duration})}{\text{(inventory holding cost per period)}}
\]

Smith’s rule is guaranteed to minimize the sum of inventory holding charges. *Inventory holding charges* for each job are computed as:

\[
\frac{(\text{inventory holding cost per period}) \times \text{(period completed)}}{}
\]

### 4.2.2 CONTROLLING TARDINESS CHARGES

The *lateness* of a job is computed as:

\[
\text{(period due)} - \text{(period completed)}
\]

A job with negative lateness is early, a job with zero lateness is exactly on time, and a job with positive lateness is tardy. A *tardiness charge* is computed only for jobs which are tardy:

\[
\text{(tardiness cost per period)} \times \text{(lateness)}
\]

Because tardiness charges apply only to jobs which are tardy, an indicator of each job’s status with respect to its deadline for completion can be useful in setting priorities to avoid tardiness. Such a quantity, called *slack*, is derived as follows:

\[
\text{(period due)} - \text{(duration)}
\]

Jobs can be prioritized in a manner analogous to Smith’s rule by considering both the likelihood that a
job will be tardy and the corresponding tardiness cost. The result, called the \textit{weighted slackness rule}, sequences jobs in increasing order of the following:

\[(\text{slack}) \div (\text{tardiness cost per period})\]

\subsection*{4.2.3 CONTROLLING CHANGEOVER CHARGES}

The costs of switching from each job to any other job in the sequence are summarized in a nonsymmetric \textit{changeover cost matrix}. The \textit{least changeover heuristic} always chooses the lowest possible changeover cost among the jobs which are not yet sequenced. The first two jobs scheduled are the two jobs with the lowest changeover cost in the matrix. The third job scheduled is the one with the lowest changeover cost when following the second job. The process of choosing the lowest feasible cost is repeated until all of the jobs have been scheduled.

\subsection*{4.3 DESCRIPTION OF THE DECISION SUPPORT SYSTEM}

With the aid of two experts in the job scheduling area, I constructed a menu-driven computer system to support the job sequencing decision described above. The original design included all of the items of information, capabilities and display formats which the expert identified as potentially relevant to the
problem. Although the design was streamlined considerably during the course of development and pretesting, it is still flexible in the sense that it provides numerous information items (far more than required by the suggested problem-solving strategy described later in this chapter). There are few rules restricting the sequence in which operators are invoked. Both the variety of information items available and the trivial rule set contribute to the system's flexibility (Silver, 1988a).

The same data sets were provided to all subjects. Each data set consists of seven jobs labelled A through G which are unrelated to jobs of the same name in other data sets. The system uses the term job set rather than data set (The terms will be used interchangeably in the ensuing discussion.). The job sets were generated from a random number table prior to the experiment. Inventory holding costs, tardiness costs and changeover costs were generated from uniform distributions whose parameters varied across job sets. In this sense, the "environment" changed for each job set. Variability in job set parameters reduces the structure of the task by making it difficult to determine the magnitude of possible cost reductions and the comparative importance of cost categories. The
parameters for the job sets used in the experiment are shown in Table 5.

I screened all of the job sets which were generated to verify that improvement over the starting cost was possible. I also attempted to establish that the required sequence modifications were not easily discovered by accident. Following the pretest, I selected the job sets which appeared to have the best discriminatory power for the experimental session. However, pretest data was limited because of the extremely small number of subjects who completed more than two job sets.

The overall structure of screens within the final version of the system is depicted in Figure 3. Each screen is described below. Reverse video is used throughout the system to identify regions of the screen containing pointers, command options, and messages (Italics is used in the figures in place of reverse video.)

An experimental session begins with a welcome message, followed by the Schedule Comparison Screen for the first data set. During the session, attempts to use a menu option inappropriately result in the display of an error message describing the problem. Keystrokes which do not correspond to a menu item result in a
beep. Once a subject has indicated and confirmed a desire to proceed to the next job set, the subject is asked to make two confidence ratings before being provided with the Schedule Comparison Screen for the next data set. Subjects are provided with paper on which to write whatever they desire, but none of the system reports can be printed. These last two interface characteristics, the confidence ratings and lack of printout capability, are discussed further in the next chapter.

4.3.1 SCHEDULE COMPARISON SCREEN

The Schedule Comparison Screen (Figure 4) provides information concerning the sequences which would be generated by the various sequencing rules described in Section 4.2. The rationale for providing this information was that the subtotals by cost category might reveal the relative room for improvement in each category. Comparison of sequences might also reveal consensus or conflict regarding the best position for a job.

After viewing the screen, the subject presses "M" to proceed to the Modification Screen. The system automatically provides the lowest total cost sequence from the Schedule Comparison Screen as the initial job.
sequence for modification. Pretesting revealed that most subjects preferred to start from the lowest total cost sequence when given a choice.

4.3.2 MODIFICATION SCREEN

The Modification Screen (Figure 5) is used to make changes to the job sequence. Three less frequently used options appear (in reverse video) along the bottom of the screen. The options which are most frequently used appear in the boxed menu at the right of the screen. In order to move a job, the subject marks the job to be moved, marks the desired location, and presses ENTER. The "Restore Previous" and "Restore Best" options allow the subject to conveniently restore an earlier job sequence. These capabilities facilitate what-if exploration of alternative sequences.

The Modification Screen also provides information concerning the computed charges for each job in the sequence by category. It is important to note that the information provided on this screen is dependent on the job sequence. In other words, the numeric items are not constant characteristics associated with each job. All of the other information available concerning the jobs in the sequence is reached via the "Analyze Schedule" option (in the boxed menu).
Both outcome and cognitive feedback are provided on the Modification Screen. After each change, the subject can compare the current grand total to the previous grand total and also to the lowest grand total of any sequence tried up to that point (best achieved). This outcome feedback is presented in an absolute rather than relative form (Jacoby, Mazursky, Kuss & Troutman, 1984): subjects know when they have reduced the total cost of the sequence but do not know how their costs compare to the costs achieved by other subjects.

Cognitive feedback is provided in the form of totals by category for current and previous charges. Subjects can use this information to determine the underlying reasons for a change in total cost. The subtotal breakdown can be used to determine why hypotheses concerning the effects of a change were confirmed or disproved.

4.3.3 ANALYSIS SCREEN

Unlike the Modification Screen, the Analysis Screen (Figure 6) includes sequence-independent job characteristics (inventory cost per period, duration of job, and tardiness cost per period tardy). The jobs are presented in the order in which they are currently
scheduled. This screen also provides entry to the remaining screens available to the user.

The feature of primary interest on this screen is a form of what-if computation called the I&T Net Effect. The user can highlight any two adjacent jobs in the sequence. The I&T Net Effect for the highlighted jobs represents the change in the sum of inventory holding and tardiness charges if the positions of the two jobs were reversed. The calculation performed takes into account all of the data in the last four columns of the display for those two jobs. A positive I&T Net Effect indicates an overall increase of the specified magnitude in inventory and tardiness charges, while a negative indicates a decrease. Of course, whether the total cost of the sequence would increase or decrease is also dependent on the impact of changeover charges, which is not included in the computation. The Net Effect calculation is explained both during the lecture on job scheduling and during training sessions. The lecture also provides a numeric example of the calculation.

The system design regarding the I&T Net Effect is intentionally inconvenient. Because the I&T Net Effect computation is central to the problem-solving strategy I developed, I am particularly interested in subjects'
use of this feature. In order to be able to monitor usage, I forced subjects to request the computation rather than simply displaying the number automatically. The Best Computed Net Effect is a convenience feature which displays the lowest net effect actually computed so far. The user can request that the system automatically highlight the two jobs associated with that number by pressing "B".

4.3.4 CHANGEOVER COSTS SCREEN

A changeover costs matrix (Figure 7) is accessible from the Analysis Screen. The matrix is organized to reflect the order in which the jobs are currently scheduled. The calculation necessary to determine the effect of a sequence change on changeover charges is not provided on this screen (but can be obtained by actually attempting the change on the Modification Screen). The correct calculation involves six numbers from the matrix (except for locations near the top and bottom of the schedule). The rationale for providing this information was that subjects could identify particularly good or bad subsequences and possibly apply a changeover heuristic to a subset of jobs.
4.3.5 HIGH/LOW VALUES SCREEN

The High/Low Values Screen (Figure 8) displays jobs with the highest and lowest rankings on a number of possible criteria. It was thought that such information might be useful in assessing the sensitivity of different criteria based on the range of values displayed. Subjects could also identify jobs which were particularly good/bad choices for the beginning or end of a sequence.

4.3.6 HELP SCREENS

Two versions of the system were actually used in the experiment, depending on the treatment condition. Both versions of the system provided screen-dependent operating help at the request of the user. Operating help explains the operation of system features as well as providing information regarding the meaning of certain items, such as I&T Net Effect, which are not intuitively obvious. The screens available through operating help are reproduced in Appendix 1.

The second version of the system provides screen-dependent strategy help in addition to operating help. Strategy help presents the strategy for using the system which I developed and concentrates on when
and how features are combined within that approach. The screens available through strategy help are presented in Appendix 2. Both types of help provided by the system are designed as a supplement to training rather than as a substitute for it. The strategy help option did not appear on any of the screens in the version of the system for subjects receiving operations training.

4.4 DESCRIPTION OF THE SUGGESTED STRATEGY

Although the system was developed with the input of domain experts, I was primarily responsible for formulating a specific problem-solving strategy to address the task. While I was not an expert in job scheduling, I was thoroughly familiar with the domain knowledge necessary to perform the task. I was also highly motivated to produce a strategy that worked and had ample time in which to do so. I was able to consult domain experts as necessary.

At the outset, we had expected that the suggested strategy would consist of a number of general rules of thumb. Contrary to expectation, however, the resulting strategy is a heuristic algorithm: The steps can be specified explicitly, but the result is not guaranteed to be optimal.
My experiences in developing the problem-solving strategy provided an invaluable frame of reference for understanding subjects' behavior. Subjects employed some of the same faulty reasoning that I had exhibited in my own early attempts at problem-solving. I also gained insight into the uses and relative strengths and weaknesses of the information available through the system. Although admittedly a biased account, the following presentation describes my strategy formulation process and the conclusions which I reached regarding information value.

Throughout the strategy formulation process, I was aware that I would have to be able to justify what I was doing in order to train others to do so. The need to constantly consider reasons and motives probably encouraged me to use a more thoughtful approach than I might have otherwise. Asking protocol subjects "to verbalize reasons and motives has substantial effects on both immediate performance and learning" (Ericsson & Simon, 1980, p. 231).

During my initial attempts to improve job sequences, I met with far more failure than success. I found the subtotals on the Modification Screen absolutely essential in understanding how my changes, both successful and unsuccessful, affected various
charges. One of my first observations was that drastic changes (moving a job several positions in a sequence) rarely met with success. I reasoned that large changes significantly altered whatever characteristics of the starting sequence were responsible for its initial advantage over the sequences derived from other rules.

Another problem with large moves was the "ripple effect": one move changed the charges associated with several jobs. I had difficulty understanding exactly what caused the changes in subtotals I observed. As a result, I avoided large moves in favor of smaller, incremental changes which were more easily understood. Restricting attention to only small moves also reduces the number of options considered at each point in the solution process. If a large move is necessary, it can still be accomplished through a series of smaller moves.

Even after restricting myself to moves of one position (first to second, sixth to fifth, etc.), I often failed to realize the desired improvements. The first problem was that I was failing to consider all of the job characteristics which determine the impact of a move within a single cost category. I discovered that the calculated charges presented on the Modification Screen could be misleading as guides for action. For
example, I was failing to properly consider differences in job durations when changing the sequence based on relative inventory holding charges. As a result, I developed a preference for the predominantly sequence-independent data provided on the Analysis Screen. I eventually arrived at the calculations necessary to correctly predict the effect of a change on a single cost category.

Through analysis of the subtotals after each change, I soon realized that by improving one cost category at a time I was failing to consider the tradeoffs which are the essence of the problem. In order to improve my performance, I would have to consider more factors simultaneously. I decided to combine consideration of inventory holding and tardiness charges because of certain similarities between them. The calculations required to predict the effect of a change are similar for both types of charges, and both of these charges vary systematically with direction: the later it is, the more it tends to cost. At this point, many hours into strategy development, the I&T Net Effect calculation was born.

I established the goal of improving sequences with respect to inventory holding and tardiness charges, and delayed consideration of changeover charges until that
subgoal had been achieved. I did not find any convenient method for accounting for changeover charges through analysis of the data on the Changeover Costs Screen. The process of using the changeover matrix is extremely slow. Six values have to be located and read from the matrix for a typical sequence change, and the numbers have to be integrated through calculation into a single number representing potential savings in changeover charges. Whatever potential savings in changeover charges are identified have to be compared against effects on inventory holding and tardiness charges.

I decided on a cognitively less demanding and faster method of accounting for changeover charges involving a systematic trial of all of the possible moves of one or two positions that might reduce changeover charges. Only small moves are attempted because improvement becomes more difficult as the size of the move increases. Larger moves are less likely to work because they are more disruptive to the compromise between inventory holding and tardiness charges already achieved. Once I had determined an appropriate pattern of moves on the Modification Screen, the strategy was complete.
The suggested strategy is organized in terms of three methods for accomplishing two subgoals. The three stages are:

1) Improve the sequence with respect to combined inventory holding and tardiness charges by using the I&T Net Effects.

2) Improve the sequence by reducing changeover costs through moves of one position.

3) Improve the sequence by reducing changeover costs through moves of two positions.

This organization is emphasized during training in order to facilitate instruction, since Kieras and Polson suggest that the user’s task representation is described by the GOMS model (1985).

Each stage uses a specific method to achieve the subgoal. The second and third stages represent two possible methods for achieving the subgoal of reduced changeover costs. My selection rule specifies that all possible moves of one position (method 1) should be explored prior to moves of two positions (method 2).

The subgoals and methods shifted as I adapted to the demands of the situation. For example, I discarded the subgoals of improving inventory holding or tardiness charges on an individual basis. I also discarded methods for reducing changeover costs which required detailed inspection of the Changeover Costs
Screen. All of the methods were refined over time, particularly after adjustments were made in the system interface (e.g., the I&T Net Effect calculation as a system option).

A narrative description of the methods is provided by the strategy help screens in Appendix 2. Summary flowcharts showing the main operations comprising each stage are presented in Figures 9 through 14. The main operator invoked for each step of the flowcharts is indicated in the bottom corner of the corresponding box/diamond. Supporting operations such as the use of arrow keys have not been depicted.

Two operators, move and user, appear in the figures but do not correspond directly to system operators. As described in Section 4.3.2, a number of separate system operators are required to move a single job to a new location. For presentation purposes, however, the separate actions are not diagrammed.

The user operator signifies a mental search or comparison when no system operator is invoked. When a system operator is indicated, user operators are assumed to exist as well. For example, a user cannot move a job without identifying its current location mentally prior to marking it. However, there is a
level of detail beyond which the presentation becomes unnecessarily cumbersome.

In my experience with the strategy, the majority of improvements are achieved in the first and second stages. The third stage is particularly conservative in the sense that it requires a fair amount of effort for what are usually small improvements. Moves of two positions are more disruptive to the compromise between inventory holding and tardiness charges than moves of one position, so they are less likely to reduce total costs. The strategy was not extended to include even more conservative moves of three or more positions. In all stages, changes are made following a steepest descent approach: the most promising alternative (high potential savings) is tested before considering other alternatives. This approach is used to avoid the backtracking which can take place when a better move is discovered only after an inferior change has been made.

The strategy entails a specific pattern of information use which reflects my views regarding information value. The goals are achieved without reference to the Schedule Comparison Screen, Changeover Costs Screen, or High/Low Values Screen. I failed to find any use for the Schedule Comparison Screen or the High/Low Values Screen. The Changeover Costs Screen is
not used because I identified a faster and less error-prone means of accounting for changeover charges.

Few of the information items (excluding menu items) presented on the two "useful" screens are essential to the strategy. For the basic form of the problem used during experimental sessions, only the Best Net Effect information is required from the Analysis Screen (Figure 6). The I&T Net Effect calculation is absolutely essential to the strategy. When considering only exchanges of adjacent jobs, the Net Effect calculation summarizes all of the other information except the changeover charges.

On the Modification Screen (Figure 5), strategy-trained subjects must refer to the column of changeover charges, the grand total, previous total and best achieved. Although I used the subtotals extensively in developing the strategy, they are not necessary in order to apply the strategy (unless the subject wishes to confirm that costs are behaving as predicted).

Although I make no claims that this strategy is the best way to solve the problem, I had several reasons to believe that it was a good way to solve the problem even before I had any data on other people's performance with the task. I felt confident in the strategy because I knew that every action could be
justified and that many alternative courses of action had been tested. I could see that my own ability to modify job sequences greatly improved as the strategy developed. I had tested the full strategy on numerous data sets and from different starting sequences within each data set. The strategy performed consistently regardless of the sequencing rule which I chose as a starting point, usually producing the same answer.

Now that I have collected performance data, I have even more reason to believe that the strategy is sound. Unlike the performance of the suggested strategy, the performance of subjects' strategies during the pretest often depended heavily on the starting order of the jobs. Even more importantly, no subject exceeded the performance of the suggested strategy on any data set.

4.5 ASSESSMENT OF THE TASK

In retrospect, it is possible to evaluate the completed task environment using the criteria for an ideal task identified in the first section of this chapter. The first criterion, an objective performance score, is easily satisfied. For this task, the performance score is the total cost of a completed job sequence.
The second criterion is more problematic, although I am satisfied that the task is better suited to my purposes than the tasks used previously. In its favor, the task represents a novel situation for subjects, and novel situations are more likely to require strategy formulation skills (Sage, 1981, p. 657; Snow & Lohman, 1984, p. 360).

The problem is in the extent to which performance scores reflected the quality of the strategies which were formulated. An acceptably wide range of performance scores was observed for most data sets, but the experimental task possessed a number of characteristics which may have diminished its power to discriminate on the basis of cost performance between decision processes of low and high quality. I underestimated the role of these characteristics because I underestimated the extent to which subjects would employ random, trial-and-error approaches even over extended periods of time.

First, the task consisted of improving a schedule which performed substantially better than a random ordering of the jobs. Hence, the initial state was closer to a low cost final state than if the subjects had started with a random ordering. Because the distance between initial and final states was reduced
relative to a random starting point, the range of cost performance scores may also have been reduced. Similarly, the likelihood of producing a low cost schedule through a series of "guesses" was increased by decreasing the complexity of the required series.

Three other task characteristics enabled subjects to produce low cost schedules even when using an essentially random guessing process. The primary characteristic facilitating a trial and error approach was the ability to try any change and reverse it without penalty. In a related vein, the number of changes which could be attempted was limited only by the subject's patience and the duration of the experiment. These two characteristics, combined with the fact that there were only seven jobs to be scheduled, created a situation where random guessing could produce low cost solutions.

It can be argued that the task environment provided subjects with a "crucial decision aid" in the form of "the iterative taking of action, monitoring of progress, and adapting of action to problem response" (Connolly, 1982). While I observed a great deal of "iterative taking of action," for most subjects I failed to see much "adapting of action to problem response." Fortunately, as discussed in the next
chapter, data concerning both process and outcomes was collected.

Regarding the final criterion, the complexity of the problem-solving strategy, this task compares favorably to the experimental tasks summarized in Table 4. Given the limited amount of time available for training (60 minutes) and practice (30 minutes), I believe that instruction in the strategy was non-trivial. Overall, the task environment provides an acceptable context for the experiment described in the next chapter.
### Table 4 (continued on next page)

**Evaluations of the Experimental Tasks Used in Previous GEFT Research**

<table>
<thead>
<tr>
<th>Reference</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Feasibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benbasat &amp; Dexter (1979, 1982)</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

**Task:** Act as inventory/production manager of simulated firm: set order points, order quantities, and daily production figures.

| Benbasat, Dexter & Masulis (1981) | + | 0 | - | +           |

**Task:** Locate price/quantity combination to maximize profit (hill-climbing problem).

| Benbasat & Dexter (1985) | + | - | - | +           |

**Task:** Allocate advertising budget across three territories to maximize profit.

**Note.** The numbered criteria correspond to: 1) objective performance score, 2) strategy formulation demands, and 3) nontrivial instruction. The symbols used to assign ratings signify the following: + for Satisfactory, 0 for Marginal, and - for Unsatisfactory.
<table>
<thead>
<tr>
<th>Reference</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Feasibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dos Santos (1982)</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td><strong>Task</strong>: Evaluate scenarios to find or prioritize potential problems in a simulated business environment.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lusk (1979)</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>Lusk &amp; Kersnick (1979)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Task</strong>: Answer questions based on information contained in a report.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pracht &amp; Courtney (1988)</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td><strong>Task</strong>: Answer questions concerning a simulated firm after observing the outcomes of numerous production and marketing decisions.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** The numbered criteria correspond to: 1) objective performance score, 2) strategy formulation demands, and 3) nontrivial instruction. The symbols used to assign ratings signify the following: + for Satisfactory, 0 for Marginal, and - for Unsatisfactory.
Table 5

Job Set Parameters

<table>
<thead>
<tr>
<th>Job Set</th>
<th>Inventory Holding Costs</th>
<th>Per Period:</th>
<th>Best Known Final Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Tardiness Costs</td>
<td>Changeover Costs</td>
</tr>
<tr>
<td>Practice Sets:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1-50</td>
<td>1-50</td>
<td>100-300</td>
</tr>
<tr>
<td>2</td>
<td>1-100</td>
<td>1-200</td>
<td>100-500</td>
</tr>
<tr>
<td>3</td>
<td>1-100</td>
<td>1-100</td>
<td>1-300</td>
</tr>
<tr>
<td>Experimental Sets:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1-100</td>
<td>1-100</td>
<td>1-300</td>
</tr>
<tr>
<td>2</td>
<td>1-50</td>
<td>1-50</td>
<td>1-300</td>
</tr>
<tr>
<td>3</td>
<td>1-50</td>
<td>1-100</td>
<td>100-300</td>
</tr>
<tr>
<td>4</td>
<td>1-100</td>
<td>1-50</td>
<td>1-200</td>
</tr>
<tr>
<td>5</td>
<td>1-100</td>
<td>1-100</td>
<td>100-400</td>
</tr>
</tbody>
</table>

Note. All data were generated from uniform distributions with the parameters indicated above. Periods of processing time required for all jobs was between 1 and 10. Period due for all jobs was between 1 and 50.
Figure 3. Structure of the decision support system screens.
**ANALYZE: Schedule Comparison**

<table>
<thead>
<tr>
<th>JOB SEQUENCE</th>
<th>Smith's Rule</th>
<th>Shortest Processing Time</th>
<th>Weighted Slackness Rule</th>
<th>Least Changeover Heuristic</th>
<th>Random Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>E</td>
<td>G</td>
<td>D</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>F</td>
<td>B</td>
<td>E</td>
<td>E</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>D</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>C</td>
<td>D</td>
<td>F</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>E</td>
<td>F</td>
<td>B</td>
<td>B</td>
<td>E</td>
<td>F</td>
</tr>
<tr>
<td>D</td>
<td>G</td>
<td>A</td>
<td>C</td>
<td>F</td>
<td>G</td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td>F</td>
<td>A</td>
<td>G</td>
<td></td>
</tr>
</tbody>
</table>

|                | Inventory $| 4663       | 5956       | 5977       | 4375       |
|                | Tardy $    | 2613       | 1437       | 608        | 944        |
|                | Changeover $| 1253       | 1073       | 739        | 567        |
|                | TOTAL $    | 7616       | 7173       | 7303       | 7488       | 7409       |

*M - Modify Schedule*  
*O - Operating Help*

*Figure 4. Schedule Comparison Screen.*
**SMITH'S RULE: Modification Screen**

<table>
<thead>
<tr>
<th>JOB</th>
<th>CHANGEOVER Charge</th>
<th>TARDY Charge</th>
<th>INVENTORY Holding Charge</th>
<th>Period COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>155</td>
<td>273</td>
<td>495</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>440</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>182</td>
<td>78</td>
<td>440</td>
<td>11</td>
</tr>
<tr>
<td>D</td>
<td>175</td>
<td>140</td>
<td>1232</td>
<td>16</td>
</tr>
<tr>
<td>G</td>
<td>210</td>
<td>0</td>
<td>770</td>
<td>22</td>
</tr>
<tr>
<td>B</td>
<td>113</td>
<td>84</td>
<td>729</td>
<td>27</td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>1080</td>
<td>396</td>
<td>36</td>
</tr>
</tbody>
</table>

**TOTALS:** 839 1655 4502 **GRAND TOTAL:** 6996

**BEST ACHIEVED:** 6996

**PREVIOUS:** 919 2003 4471 **PREV. TOTAL:** 7393

- **F** - Finished Making Changes
- **O** - Operating Help
- **S** - Strategy Help

Figure 5. Modification Screen.
**ANALYZE: Schedule Data**

<table>
<thead>
<tr>
<th>JOB</th>
<th>CHANGEOVER Cost</th>
<th>INVENTORY Cost per Period</th>
<th>DURATION of Job</th>
<th>TARDINESS Cost per Period</th>
<th>TARDINESS Tardy</th>
<th>Periods LATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>293</td>
<td>41</td>
<td>6</td>
<td>39</td>
<td></td>
<td>-34</td>
</tr>
<tr>
<td>F</td>
<td>39</td>
<td>39</td>
<td>9</td>
<td>16</td>
<td></td>
<td>-15</td>
</tr>
<tr>
<td>C</td>
<td>279</td>
<td>38</td>
<td>9</td>
<td>34</td>
<td></td>
<td>-2</td>
</tr>
<tr>
<td>A</td>
<td>274</td>
<td>34</td>
<td>10</td>
<td>16</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>E</td>
<td>166</td>
<td>9</td>
<td>3</td>
<td>42</td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>D</td>
<td>130</td>
<td>6</td>
<td>6</td>
<td>50</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>G</td>
<td>111</td>
<td>5</td>
<td>9</td>
<td>43</td>
<td></td>
<td>33</td>
</tr>
</tbody>
</table>

I&T Net Effect: 

Best Computed Net Effect:

**B - Best Net Effect (Highlighted)**

**C - Compare Schedules**

**E - Evaluate Net Effect**

**H - Hi/Low Values**

**T - Changeover Table**

**M - Return to Modification Screen**

**O - Operating Help**

*Figure 6. Analysis Screen.*
**ANALYZE: Changeover Costs**

Job which FOLLOWS:

<table>
<thead>
<tr>
<th>JOB</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>160</td>
<td>----</td>
<td>71</td>
<td>6</td>
<td>229</td>
<td>293</td>
<td>120</td>
</tr>
<tr>
<td>F</td>
<td>139</td>
<td>83</td>
<td>279</td>
<td>238</td>
<td>146</td>
<td>----</td>
<td>234</td>
</tr>
<tr>
<td>C</td>
<td>274</td>
<td>89</td>
<td>----</td>
<td>33</td>
<td>295</td>
<td>296</td>
<td>135</td>
</tr>
<tr>
<td>A</td>
<td>----</td>
<td>220</td>
<td>206</td>
<td>33</td>
<td>166</td>
<td>92</td>
<td>93</td>
</tr>
<tr>
<td>E</td>
<td>265</td>
<td>215</td>
<td>180</td>
<td>130</td>
<td>----</td>
<td>89</td>
<td>81</td>
</tr>
<tr>
<td>D</td>
<td>175</td>
<td>38</td>
<td>123</td>
<td>----</td>
<td>4</td>
<td>116</td>
<td>111</td>
</tr>
<tr>
<td>G</td>
<td>199</td>
<td>185</td>
<td>72</td>
<td>61</td>
<td>236</td>
<td>54</td>
<td>----</td>
</tr>
</tbody>
</table>

Press any key to continue...

*Figure 7. Changeover Costs Screen.*
### ANALYZE: High/Low Values

<table>
<thead>
<tr>
<th></th>
<th>INVENTORY HOLDING COST</th>
<th>TARDY COST</th>
<th>SLACK TIME</th>
<th>PROCESSING TIME</th>
<th>CHANGEOVER COST</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HIGHEST:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>41</td>
<td>D</td>
<td>50</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>F</td>
<td>39</td>
<td>G</td>
<td>43</td>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td><strong>LOWEST:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>6</td>
<td>C</td>
<td>34</td>
<td>E</td>
<td>B</td>
</tr>
<tr>
<td>G</td>
<td>5</td>
<td>A F</td>
<td>16</td>
<td>G</td>
<td>E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>SMITH's RULE RATIO</th>
<th>WEIGHTED SLACKNESS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Processing Time ÷</td>
<td>Slack ÷ Tardy Cost</td>
</tr>
<tr>
<td></td>
<td>Inventory Holding Cost</td>
<td></td>
</tr>
<tr>
<td><strong>FIRST:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.15</td>
<td>G 0.23</td>
</tr>
<tr>
<td>F</td>
<td>0.23</td>
<td>E 0.31</td>
</tr>
<tr>
<td><strong>LAST:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>1.00</td>
<td>A 1.06</td>
</tr>
<tr>
<td>G</td>
<td>1.80</td>
<td>F 1.31</td>
</tr>
</tbody>
</table>

Press any key to continue...

*Figure 8. High/Low Values Screen.*
Figure 9. Flowchart of Stage I of suggested strategy.
Figure 10. Flowchart of Stage II of suggested strategy.
Figure 11. Changeover alternatives considered in Stage II.
Figure 12. Flowchart of Stage III of suggested strategy.
Figure 13. "Up" alternatives considered in Stage III.
Figure 14. "Down" alternatives considered in Stage III.
Chapter 5. EXPERIMENTAL DESIGN

A laboratory experiment was designed to explore the research questions presented in Chapter 3. Both main and interaction effects of two variables were of interest: The first dimension was cognitive restructuring ability, and the second dimension was the type of guidance provided through training and on-line help.

5.1 THE PRETEST

The pretest subjects were MBA students in an introductory operations management course. The students were asked to take part in the study in exchange for being exempted from one problem on the midterm exam. Three students declined to participate. The 29 students who consented to participate were administered the Group Embedded Figures Test (Witkin et al., 1971) prior to the system training sessions. Students were matched as closely as possible on the basis of their GEFT scores and then alternately assigned to the training treatments. A listing of group assignments and locations was distributed to the students prior to the training sessions.
A lecture on job scheduling was presented one week prior to the system training sessions. Although I was not aware of this at the time, the instructor informed the students prior to the lecture that the material would not be tested and that their performance on the experimental task would not affect their grade in any way. One of the subjects informed me of these circumstances during a discussion near the close of pretesting. I have no doubt that the instructor’s comments adversely affected both the students’ knowledge of job scheduling and their motivation.

For training purposes, I was given access to the subjects during the regularly scheduled class meeting time for two hours. Operations training was conducted during the first hour, and strategy training was conducted during the second hour. The subjects were informed during the training sessions that a practice version of the system was available to them in the computer lab, but only four subjects chose to use the system between the training sessions and the individual sessions (as indicated by a sign-in sheet which lab staff presented prior to releasing the system disks).

Following the training sessions, the subjects scheduled individual appointments with me to use the system. The subjects used the system for one hour or
until they had exhausted the data sets, whichever came first. The subjects then completed a taped exit interview which required less than fifteen minutes. The exit interview requested retrospective information concerning strategies, system use, perceptions of the system, and the extent to which the subject had communicated with others concerning the solution of the problem. The individual sessions were conducted over a period of approximately one month, so the delay between training and use was anywhere from two days to 27 days.

In order to counter some of the variation between subjects in the delay between training and use, strategy-trained subjects were provided with a review of the suggested strategy in text form (essentially the same information as provided by the system's strategy help feature). Subjects were presented with the review text, asked to read it, and informed that it would not be available once they had begun to use the system.

All subjects were informed that performance was determined by the quality of solutions and not the speed with which the solutions were obtained. All subjects were instructed to "talk aloud" during the entire session and completed two "talk aloud" warm-up exercises (based on Ericsson & Simon, 1984, pp. 375-
379). All subjects were also given an opportunity to ask me questions prior to beginning, and informed that they would be unable to ask me questions once they had begun. I remained present during the entire session (sitting behind the subject), took notes concerning the behavior I observed, and tape recorded the verbal protocols.

Pretest results in terms of speed and performance will be discussed later in this chapter. In addition to providing evidence in support of some hypotheses, pretesting served a number of purposes:

1) The pretest demonstrated that more than an hour would have to be allowed for training and practice.

Operations training required about 45 minutes, leaving only 15 minutes for subjects to practice. Strategy training required almost the full hour, leaving virtually no time for practice. In both cases, the training time required was about 15 minutes beyond my expectations.
2) The pretest established that the suggested strategy could be understood and used by at least some of the subjects in the limited amount of time available for training.

Given a very real practical constraint on the amount of time available for training, it was important to establish that strategy training would be feasible.

3) The pretest revealed that some data sets did not create a sufficient range of performances.

The data sets which were demonstrably poor discriminators were not used for the subsequent experiment. Unfortunately, two factors limited the information available concerning the discriminatory power of data sets. First, as described above, I had reason to doubt both the subjects' knowledge and their motivation. Their performance and behavior might have been substantially different from the performance of a more knowledgeable, better-motivated group. The operations-trained subjects used in the actual experiment did have better average cost performance for all comparable job sets, suggesting that pretest data set performance was not a perfect indicator of data set performance during the actual experiment.

Second, the design of the pretest system limited the amount of data collected. The pretest system allowed subjects to choose and modify any or all of the
five job sequences presented on the Schedule Comparison Screen. Subjects who chose the very thorough approach of modifying all five sequences could require up to an hour to produce a single performance measure for a job set. Virtually no information was available concerning later job sets because few subjects reached them. In general, the pretest proved more useful in deciding which data sets to discard than in deciding which data sets to retain.

4) The system was modified in accordance with observations made during pretesting. The most fundamental design change was to allow schedule modification to proceed from only one starting sequence. One reason for this change was to increase the number of performance measures obtained for a subject during an experimental session. This modification was also made in order to reduce the chances of "accidentally" discovering low cost schedules by making large numbers of modifications from different starting points. The final reason for making this change was to reduce the complexity of the interaction. Several subjects were confused regarding the distinction between alternative starting sequences and alternative job (data) sets. The corresponding system operators were frequently confused.
Two minor modifications affecting the decision support system involved correcting a system bug related to file management and changing one menu item from a letter to ENTER. Subjects invariably tried to use ENTER to indicate that they wished to calculate the costs of a modified schedule. Hence, their expressed preference was incorporated into the final design.

5) The format of items used to measure perception/confidence were modified due to subjects' expressed confusion and/or interpretation of those items.

Because of the talk aloud instructions and the oral administration of the exit questionnaire, subjects provided information concerning item interpretation which might never have been obtained under different conditions. For example, many of the subjects expressed difficulty in understanding two of the items taken directly from Larcker and Lessig's (1980) instrument for measuring the perceived usefulness of information. In order to alleviate the confusion observed during the pretest, the wording of the items was altered slightly for the experimental sessions, and one of the original items was omitted.

After each job set, subjects were asked to indicate on a scale from 0 to 7 how certain they were that they had "found the lowest cost schedule possible
for the job set...just completed." While the scales and wording differ slightly, this item format is quite common in MIS research (e.g., Schroeder & Benbasat, 1975; McGhee, Shields & Birnberg, 1978; Gul, 1984). The confidence ratings subjects provided seemed to vary depending upon the standard which the subject used in responding. Some subjects seemed to focus on the cost they had obtained, while others mentioned their decision process or domain knowledge. As a result, I decided that separate confidence items for outcome and process were in order. Subjects also seemed to vary in whether they considered the "best" cost to mean the lowest cost obtained through exhaustive search, or the "best" cost obtained through "reasonable" but limited search. As a result, I decided to design the confidence items for the subsequent experiment with the intention that all subjects would use the same standard of reasonable search in responding to the items.

6) The pretest revealed potential difficulties in obtaining verbal protocol data. Even though all of the subjects verbalized quite fluently during the warm-up exercises, the rate of verbalization during system use was markedly reduced (despite repeated reminders to keep talking). Because the majority of the subjects had not practiced using
the system beforehand, I attribute most of their
difficulty in verbalizing to the high cognitive demands
of the task. Ericsson & Simon (1980, p. 237) observed
that "subjects tend to stop verbalizing or to verbalize
incompletely in conditions in which they are giving
indications of being under a high cognitive load." I
attempted to resolve this problem by ensuring that the
experimental subjects would have adequate time to use
the system prior to verbalizing. I also made the warm-
up exercise for the experiment more difficult so that
it would more accurately reflect the conditions under
which the subjects would later verbalize.

Even when subjects did verbalize, they often spoke
so softly that the taped verbalizations were not
audible over the fan noise from the computer. This
problem was partly solved by acquiring a small, clip-on
microphone. The subjects may have mumbled because they
were concerned about the impression they were making on
me, particularly when they were uncertain about their
domain knowledge and/or decision process (Biehal &
Chakravarti, 1989, pp. 84-85). Additional practice and
more careful lecturing on job scheduling were expected
to improve user confidence.
5.2 THE EXPERIMENT

Usable experimental data was obtained from a total of 42 part-time MBA students. The GEFT was administered at the time the subjects agreed to participate in the experiment. Students were matched as closely as possible on the basis of their GEFT scores and then alternately assigned to the training treatments. A listing of group assignments and locations was distributed to the students prior to the training sessions. The number of subjects in each treatment-by-ability group is shown in Table 6.

The subjects were enrolled in three different sections of an introductory operations management course within the same college and taught by the same instructor using the same notes. The lectures and training sessions were administered separately to each course section during the regularly scheduled class meeting time. For two of the sections, a lecture on job scheduling took place one week prior to the first system training session. For the third section, the lecture was delivered two weeks prior to the first system training session (The added delay was due to spring break.). Each of the lectures required approximately two hours and 30 minutes.
In order to assure that the students had the theoretical background necessary to complete the experimental task, I had developed the notes and examples concerning job scheduling in conjunction with the instructor. This guaranteed that each item of information that was provided to the subjects during the experiment and all of the essential relationships and tradeoffs had been explained during a lecture prior to the experimental session. In particular, the net effect calculation was explained both during the lecture on job scheduling and again during training.

Subjects were asked to participate on a voluntary basis but were informed that course credit would be awarded to the highest performers in each course section (Alternative means of obtaining extra credit were available upon request.). Only one student did not attend the experimental session. Prior to the experimental session, subjects were also informed that they would receive confidential reports of their performance on the experimental task and a complete debriefing.

The extra credit was considered important since Witkin et al. (1977, p. 20) cite evidence suggesting that the form of reinforcement may interact with field independence. While field independence and CRA are not
equivalent, they are related (see Section 2.3.1). External rewards are believed to motivate both field independents and field dependents (Witkin et al., 1977, p. 20).

For this experiment, the use of part-time MBA students as surrogates for managers seems justifiable. Basic cognitive abilities (such as CRA) "should not vary significantly among MBA students and managers" (Jarvenpaa, 1989, p. 298). All of the subjects had been exposed to the required facts and relationships during the lecture on job scheduling. As described in the previous chapter, the knowledge base required for the task is neither large nor complex. Extensive business experience would not appear to provide any particular advantage for this specialized task.

5.2.1 THE EXPERIMENTAL SESSIONS

I was given access to each course section during the regularly scheduled class time for a total of six hours. Three hours were allocated to each treatment group. For all three sections, experimental sessions for operations-trained subjects were conducted one week prior to the sessions for strategy-trained subjects. This sequence was established to preclude communication of the suggested strategy to operations-trained
subjects. This also assured that any advantage in terms of recall of lecture material would be with the operations-trained subjects.

For all subjects, the experimental session consisted of a system training period, 30 minutes of hands-on practice, an hour of "actual" use where subjects competed for the extra-credit prizes, and 30 minutes allocated to completion of an exit questionnaire. In practice, the sessions could not begin until all of the subjects had arrived, leaving less than 30 minutes for completion of the questionnaire. Subjects were allowed to ask questions during the training and practice periods, but not during the hour of competitive use. At the beginning of the session, all subjects received an envelope containing labelled packets of paper (separate sheets for notes during training, practice, and actual use), the exit questionnaire, and a symbolic gesture of appreciation: a Hershey's chocolate bar.

During training, a projection unit displayed the computer screen as I demonstrated the system. The training format (oral instruction followed by hands-on practice) was designed to provide clarification of the training material through demonstration, response to user questions, and access to the actual problem
environment. These are three of the forms of environmental feedback identified by Simon & Hayes (1976, pp. 276-277) which simplify the task of understanding complex instructions. I personally conducted all of the training sessions.

Because of the capacity limits of short-term memory, help features were incorporated into the system so that subjects could review the information presented during training as often as necessary (Simon & Hayes, 1976, pp. 274-275). Subjects were informed at the beginning of the training session that, while they were free to take notes, the system's help operator included essentially the same information as I was presenting. The help operator was demonstrated before any other operator. On-line help, as opposed to a summary sheet, was used so that actual help usage could be monitored.

The operations-trained subjects received approximately 45 minutes of training (The exact length depended on the nature and number of questions which subjects asked.). No strategic content was included in the training which consisted solely of a demonstration of all of the system's capabilities. In Silver's terminology, the training consisted of informative mechanical guidance (1989). A partial review of job scheduling concepts was provided along with the
discussion of the corresponding information items available on the screens.

The strategy-trained subjects received approximately 60 minutes of training which integrated both operating and strategic information, what Silver would call suggestive decisional guidance for both structuring and executing processes (1989). Strategy training presented a superset of the information provided during operations training. Besides being told how each of the system operators fit into an overall problem-solving approach, the subjects were told why and how the approach worked. Strategy-trained subjects had access to both operating and strategy help. Strategy help summarizes the integrated information provided during the training session in an on-line, written form.

Subjects used the decision support system described in the previous chapter to modify up to five job sets. They were provided with paper on which to write whatever they desired, but none of the system reports could be printed (in order to monitor screen usage). All subjects completed the same job sets in the same order. Because data for all subjects was desired for as many iterations as possible, the order of the data sets was not randomized. Subjects were
instructed explicitly that they should emphasize cost performance over speed. This priority was reinforced by explaining that only average performance would be used to determine the winners of extra credit. Hence, no advantage would be obtained by working quickly.

5.2.2 VERBAL PROTOCOL SESSIONS

Students in the third course section were offered unconditional course credit in exchange for participating in an additional session following the experimental session (Alternative means of obtaining extra credit were available upon request.). All of the students agreed to participate in the additional session and scheduled individual meeting times with me. The delays between the experimental session and the individual sessions ranged from two days to nine days. In all cases the individual sessions were completed prior to the class debriefing.

The first purpose of the individual sessions was to obtain taped concurrent verbal protocols during system use. Subjects were given "talk aloud" instructions (based on Ericsson & Simon, 1984, pp. 375-379) and completed a "talk aloud" warm-up exercise. All subjects were given an opportunity to ask me questions prior to beginning, and informed that they
would be unable to ask me questions once they had begun. I remained present during the entire session (sitting behind the subject), took notes concerning the behavior I observed, and tape recorded the verbal protocols. The subjects took as long as they liked to complete two job sets and verbalized throughout the session. Following completion of the job sets, I asked some subjects specific questions concerning their strategies. Twelve usable protocols were obtained.

The second purpose of the individual sessions was to explore the issues of maintenance and generalization of training. In order to compare delayed performance with performance on the day of training, the first job set for the individual sessions corresponded to the third job set from the experimental session. After the subjects had furnished confidence ratings for the first job set, I informed them that the rules for the second job set were slightly altered. They were told of an additional constraint on the solution to the second job set: Job A could not be late. A programming modification made it impossible for the subjects to leave the job set if the constraint was not met. The second job set was used to explore whether training would generalize to the new task environment.
5.3 DATA COLLECTED FOR ANALYSIS

Data was collected using a number of methods. The need for multimethod approaches in studying decision processes "is generally acknowledged" (Einhorn & Hogarth, 1982). Because the data collection techniques possess different sets of advantages and disadvantages, each method can be used to complement the data provided through another.

5.3.1 THE INDIVIDUAL DIFFERENCE MEASURE (GEFT)

The Group Embedded Figures Test (Witkin et al., 1971) was used to measure cognitive restructuring ability. In addition to the conceptual reasons for this choice, the reliability and validity of the test instrument are well-documented (Witkin et al., 1971; Desanctis & Dunikoski, 1983; Kepner & Neimark, 1984; Moran, 1983; Thompson & Melancon, 1987).

5.3.2 USER LOG

Besides automatically capturing time and performance data, the system creates an exhaustive log of user keystroke activity. Users were not permitted to print the screens so that their activity could be more closely monitored. There would also have been
practical problems associated with allowing the users to print screens, since not all computers were connected to printers. Users were allowed to take notes in order to minimize the potential impact of the printout restriction on their decision processes.

The computer log also records user responses to two confidence items which follow each job set. The confidence items are shown in Figure 15. The first item concerns confidence in the performance outcome, and the second item concerns confidence in the decision process. In order to encourage all subjects to use the same standard of "reasonable search," they are instructed to respond with respect to "schedules which can be created after attempting 100 or fewer changes."

As discussed above (see Section 5.1), the item format was based on observations of response behavior during the pretest.

The advantage of the user log is that it provides accurate and objective data without disrupting the decision process. The weakness of this approach is that it "provides no specific insight into an individual's evaluation or assessment of the available information" (Todd & Benbasat, 1987, p. 496). While the user log indicates which screen a subject requested, it cannot indicate which information items
were used (with the exception of net effects). The data collection techniques described below, exit questionnaires and verbal protocols, were used to provide insight into the nature of users' decision processes.

5.3.3 EXIT QUESTIONNAIRE

The exit questionnaire was the same for all subjects with the exception of the first page which differed by treatment. The two alternative forms of the first page and the remaining pages appear in Appendix 3. The questionnaires used during the experiment were double-sided so that examples of each screen appeared opposite the page requesting information concerning that screen (The printouts were equivalent to Figures 4 through 8.). Subjects responded to the questionnaire in writing. The majority of the questionnaire items request retrospective information concerning the nature of the decision process, the information items used, and the way information items were incorporated into the decision.

The exit questionnaire contains two sets of items designed to measure differences in perceptions of the system. The first set of questions are the open-ended
probes for each screen of information to determine if
the system failed to provide desired capabilities,
either through problems with the information actually
presented or through failure to provide desired
information. The last two pages of the questionnaire
(items A, B and C on page 7; items A and B on page 8)
measure perceptions of the usefulness of the
information provided by the system as a whole. The
items were adapted from Larcker and Lessig (1980).

Item D on page seven was included to determine
whether there were differences in the perceived
difficulty of formulating a strategy for using the
information presented. Item E on page seven was
included as a measure of users' overall satisfaction
with their strategies. Item C on page 8 was included
to measure user perceptions of the adequacy of their
own domain knowledge.

While the exit questionnaire provides summary data
concerning user processes, the reports were provided in
a retrospective, written fashion. Retrospective
reports may be inaccurate due to memory distortion,
interpretation, and/or forgetting (Todd & Benbasat,
1987, p. 497). The written format has been found "to
be linked to evaluation and censorship" (Ericsson &
Simon, 1980, p. 241). The final method of data
collection, concurrent verbal protocols, was used to supplement the data provided by the questionnaires.

5.3.4 VERBAL PROTOCOLS

Many authors have called for the use of protocol analysis in "supporting or explaining empirical results based on another methodology" (e.g., Jarvenpaa, 1989, p. 300; Lewin, 1982, p. 315; Todd & Benbasat, 1987). The circumstances under which protocols are collected can have "a significant effect on what is verbalized and on the interpretation of the verbal data" (Ericsson & Simon, 1980, p. 218). The subjects in this study reported what they were doing as they were doing it (concurrent verbalization). They were given "talk aloud" instructions (based on Ericsson & Simon, 1984, pp. 375-379) and were at no time asked to explain their reasons or motivations (although brief exchanges between the subject and experimenter did take place, usually initiated by a subject's question). The task lends itself to verbal reporting. Under these conditions, verbalization was expected to have only minor impacts, if any, on the course of the decision process or the speed of task performance (Ericsson & Simon, 1980).
There is some concern that the presence of an experimenter during verbalization may be sufficient to change the course of the decision process:

Subjects may adopt whatever strategy they think the experimenter expects, or a strategy designed to make them look less the fool and more the careful, logical decision maker (Lichtenstein, 1982, pp. 83-84).

In order to minimize the distraction due to my presence, I sat behind the subject (as suggested by Ericsson & Simon, 1984, p. 375).

While I gathered from the language used during the protocol sessions that the subjects had not forgotten that I was there, I did not get the impression that the subjects were uncomfortable. Any initial nervousness or stiffness seemed to disappear as they became more involved with the problem. In a study by Jarvenpaa (1989), 20% of the pretest participants complained that the experimenter's presence adversely affected their performance. However, "the experimenter's presence or absence was found to have no effect on the acquisition and evaluation directions in the...study" (Jarvenpaa, 1989, p. 291).

Unlike the pretest subjects described in Section 5.1, the experimental subjects verbalized audibly and fluently. They did not appear embarrassed about the
quality of their decision processes or make excuses, even when excuses appeared to me to be in order. For example, a subject with one of the more puzzling decision processes asked me "Have you had any bizarre reasonings? Anything really bizarre?" The computer completed its calculation, and she resumed verbalizing without waiting for an answer.

5.4 HYPOTHESES

The research hypotheses are grouped according to the component of the research framework being addressed. Hypotheses are stated in the alternative form where prior findings provide a basis for predicting differences between groups. Otherwise, null hypotheses are not stated formally. However, the analysis will include tests of null hypotheses concerning main and interaction effects for all dependent variables.

5.4.1 COST PERFORMANCE OUTCOMES

Cost performance data was recorded automatically in the user log files. The hypotheses regarding cost performance were presented informally in Chapter 3. The expected relationship between CRA and cost performance for each treatment is depicted in Figure
16. The relationship depicted in Figure 16 translates into the following alternative hypotheses:

**Main effect of the training treatment:**

H1(p): Strategy (S) trained subjects will have better cost performance than operations (O) trained subjects:
(S > O).

The hypothesized main effect of the training treatment accounts for the expected ordinal interaction. This prediction is based on the assumptions that the suggested strategy is superior to independently devised strategies and that users are able to adopt it. These assumptions were supported by pretest results. It should be noted, however, that the pretest subjects appeared to have inadequate knowledge of job scheduling and low motivation. The observed differences between the treatment groups in the pretest may have been exaggerated because the operations-trained subjects lacked the basis and motivation for devising good strategies of their own.

A main effect of CRA is not predicted because high CRA is expected to result in better performance only within the operations training treatment, as indicated by the hypothesized interaction:
Interaction effect of training treatment with CRA:

H2(p): Low (L) analytic, strategy (S) trained subjects will perform equally with high (H) analytic, strategy (S) trained subjects: (LS = HS).

High (H) analytic, operations (O) trained subjects will perform better than low (L) analytic, operations (O) trained subjects: (HO > LO).

High (H) analytic, strategy (S) trained subjects will perform at least as well as high (H) analytic, operations (O) trained subjects: (HS >= HO).

I propose that low analytic subjects exhibit lower performance primarily because of failures in strategy formulation rather than failures in strategy execution. In the strategy training condition, the subjects are not required to develop their own processes. Hence, task demands on CRA are greatly reduced and the primary basis for performance differences has been eliminated.

For the operations-trained subjects, demands on CRA are high. Given the analytical task used in this experiment, high quality strategies are necessarily analytical in nature while low quality strategies rely on random guessing. As discussed in Section 2.3.3, high analytics outperform low analytics in such settings (when a performance difference exists). A performance difference between operations-trained high and low analytics was not supported by pretest results, but the lack of adequate domain knowledge could well
have overshadowed any differences due to cognitive restructuring ability. As Ericsson and Simon observe, "if someone lacks the knowledge necessary for performing a task, no number of general skills will save the day for him" (1984, p. 215). Some of the pretest data sets also had severely restricted performance ranges.

5.4.2 TIME PERFORMANCE OUTCOMES

Time performance data was recorded automatically in the user log files. No specific predictions regarding time performance can be made.

The expectation is that there will be a difference in time performance for the treatments, but no directional prediction is made because of the impossibility of predicting the form of user strategies in the operations training condition. The added complexity of developing a good strategy might result in longer required time for operations-trained subjects. On the other hand, if simple processes are developed, the operations-trained subjects might complete schedules more quickly than the strategy-trained subjects.

There may also be time performance differences associated with CRA. The literature, however, contains
conflicting indications as to what the direction of the difference might be. For example, Benbasat & Dexter (1979) found that low analytics required more time to make decisions than high analytics. In contrast, Benbasat & Dexter (1982) found that high analytics required more time than low analytics to make decisions with a decision aid.

It is impossible to predict the nature of an interaction without a better understanding of the role of CRA in both strategy formulation and strategy execution. In the case of strategy-trained subjects, the difference between high and low analytics might be comparatively minor if there is no significant difference in their strategy execution skills and the individuals all use the suggested process. For operations-trained subjects, there will probably be a wide variety in the strategies employed and in the corresponding times to solution.

5.4.3 THE STRATEGY FORMULATION PROCESS

A user's strategy reflects the strategy formulation process through which it was created. An important measure, then, of the strategy formulation process is the nature of the product which is produced. Responses to the exit questionnaires and the user logs
provide data concerning the information requirements of user strategies and the procedures subjects employed. The concurrent verbal protocols indicate the nature of user strategies and sometimes provide a direct glimpse into the strategy formulation process as it is occurring.

It is an experimentally undesirable characteristic of tasks, including the one used in this experiment, that "the same outcome may be generated through different problem-solving approaches or strategies" (Todd & Benbasat, 1987). Section 4.5 discussed the characteristics of the experimental task which may have diminished its power to discriminate on the basis of cost performance between decision processes of low and high quality. However, the predictions concerning cost performance are based on the assumption that outcomes reflect differences in the processes employed by subjects. For this reason, the questionnaire and protocol data will be used to establish that qualitative differences in decision processes explain observed differences in outcomes (or exist despite the lack of marked differences in performance scores).
5.4.3.1 INFORMATION USE

The user logs can be analyzed to determine the frequency with which the net effect and best net effect operators were used. The questionnaires provide data concerning subjects' self-reported use of other specific information items. The expected pattern of performance differences is assumed to depend on process differences, giving rise to the predictions presented below.

Main effect of the training treatment:

H1(i): The information use of strategy trained subjects will differ from the information use of operations trained subjects: \( S \not\equiv O \).

The pattern of information use required by the suggested strategy is known. A greater proportion of operations-trained subjects are expected to use information items which are not required by the suggested strategy. With the exception of information items which are essential to task completion (e.g., total cost information), a greater proportion of strategy-trained subjects are expected to use the information items associated with the suggested strategy. I base these predictions on the assumptions that: 1) Strategy-trained subjects will adopt the
suggested strategy, and 2) The mere availability of items will be sufficient to encourage some use among the operations-trained subjects.

The frequency of use of the net effect calculation is also expected to differ by treatment. The net effect calculation is essential to the suggested strategy, so strategy trained subjects are expected to use it more frequently than operations trained subjects. Based on the information requirements of the suggested strategy (see Section 4.4), the specific pattern of results expected with regard to information items is summarized in Table 7.

Interaction effect of training treatment with CRA:

H2(i): The information use of low analytic, strategy-trained subjects will not differ from the information use of high analytic, strategy-trained subjects:

\( (LS = HS) \).

The information use of high analytic, operations-trained subjects will differ from the information use of low analytic, operations trained subjects:

\( (LO \not< NO) \).

Within the strategy training condition, no differences in information use are expected between high and low analytic subjects so long as they adopt the suggested strategy. Performance is expected to be the same because the process employed by both groups (and the
information required by that process) is expected to be the same regardless of CRA.

For the operations training condition, the differences in information use related to cognitive restructuring ability are expected to reflect the greater ability of high analytics to isolate relevant information (see Section 2.3.1). Judgements regarding information value in this context require an understanding of and attention to causal relationships within the data. Attention to causal relationships within the data is more characteristic of a high analytic decision process (Clark & Roof, 1988). The specific predictions presented in Table 8 focus on the information items which are most useful in understanding causal relationships (in my own, admittedly biased, experience).

The subtotal information provided on the Modification Screen is of use in understanding total cost behavior because it allows the user to see the effects of changes on each category of cost. The subtotals clarify relationships in the data, so high analytics within the operations training treatment are expected to make greater use of subtotal information.
The net effect calculation on the Analysis Screen is the only means provided by the system to integrate information without actually moving jobs. The net effect also provides information concerning causal relationships because it helps the user predict the effect of a move. Again, high analytics are expected to make greater use of the net effect (and best net effect) information. This prediction mirrors the expected difference between strategy-trained and operations-trained subjects. Judgments concerning information relevance can be facilitated by training or higher ability. In this case, operations-trained high analytics are expected to discover for themselves what would have been conveyed explicitly during strategy training.

The Changeover Costs Screen is not used in the suggested strategy because of cost/benefit tradeoffs with respect to strategy execution. The screen can be used with appropriate calculations, however, as part of an analytical process to understand the impact of a move on changeover costs. Once again, operations-trained high analytics are expected to use the Changeover Costs Screen more often.

Finally, my experience suggests that most of the data provided by the system is unnecessary or
irrelevant, although all of it is clearly related to the problem. Because of their greater ability in separating relevant from irrelevant information, I expect high analytics to use fewer information items overall.

5.4.3.2 USER STRATEGIES

Data concerning the nature of user strategies is available from exit questionnaire responses and the concurrent verbal protocols. Again, the pattern of expected performance differences will presumably be accompanied by an equivalent pattern of strategy differences.

Main effect of the training treatment:

H1(s): The strategies of strategy trained subjects will differ from the strategies of operations trained subjects: (S <> O).

The strategies of strategy-trained subjects should more closely resemble the suggested strategy, and the way in which the net effect calculation is incorporated into the decision should more closely resemble the suggested strategy for strategy-trained subjects. Verbal protocol data will be examined to determine if there are any differences in behavior with respect to specific categories of verbalizations (the protocol
coding scheme will be discussed in detail in the
analysis chapter). The number of alternative solutions
considered (as determined from the user logs) will also
be contrasted by treatment.

*Interaction effect of training treatment with CRA:*

\[ H_2(s): \text{ The strategies of low analytic, strategy-}
\text{trained subjects will not differ from the}
\text{strategies of high analytic, strategy-trained}
\text{subjects: (}LS = HS\text{).} \]

The strategies of low analytic, operations-
trained subjects will differ from the
strategies of high analytic, operations trained
subjects. \((LO \not< HO).\)

For strategy-trained subjects, the first page of the
exit questionnaire includes an item measuring the
perceived extent to which the suggested strategy was
followed. On the basis of this item, questionnaire
responses, user logs and the verbal protocols, low
analytics are expected to be indistinguishable from
high analytics.

For operations-trained subjects, the questionnaire
responses, user logs and verbal protocols are expected
to differ depending on CRA. First, the decision
processes of high analytics should be more "analytic"
than the decision processes of low analytics.
Operational descriptions of high and low analytic
processes were developed after consulting a number of
sources (Huysmans, 1970; Doktor & Hamilton, 1973; Zmud, 1978; Benbasat & Dexter, 1979; Rouse & Rouse, 1982; Clark & Roof, 1988; Hunt, Krystofiak, Meindl & Yousry, 1989). The descriptions appear in Table 9. The second area of differences to be investigated is the way in which the net effect calculation is used. High analytics should use the net effect calculation in a manner more closely resembling the suggested strategy than the way in which it is used by low analytics. Third, high analytics should produce more verbalizations associated with analytic behavior (the protocol coding scheme will be discussed in detail in the analysis chapter). Finally, the number of alternative solutions explored by each ability group will be investigated.

5.4.4 THE STRATEGY EXECUTION PROCESS

It is often difficult or impossible to distinguish between events that reflect failures in the strategy formulation process and events that reflect failures in the strategy execution process. For example, it is often impossible to distinguish between a problem caused by forgetting (The strategy was formulated properly but there has been a memory retrieval failure.) and a problem caused by inadequate strategy
formulation (Incomplete guidelines or an encoding failure resulted in execution difficulty.). As another example, a keystroke error could reflect an incorrectly encoded instruction concerning operator use (strategy formulation) or an erroneous translation of the instruction into action (a typographical error--strategy execution). As a simplifying assumption, all of the data reflecting user errors, difficulty or uncertainty concerning appropriate actions will be interpreted as measures of the strategy execution process for purposes of stating and testing hypotheses.

The training treatments were designed to have different impacts on the strategy formulation process, but I have no expectations concerning the existence of strategy execution effects related to training. Error data and help usage data from the user logs and selected coding categories from the protocol analysis will be used to explore the relationship between training method and strategy execution.

I have no basis for predicting the nature or existence of an interaction effect, but there is some evidence that there may be a main effect of CRA on the strategy execution process, as stated below.
Main effect of cognitive restructuring ability:

H1(e): Low analytics will differ from high analytics with respect to measures of the strategy execution process: (L ↔ H).

Reardon et al (1982) found evidence of memory performance differences even when high and low analytics employed the same strategy. Another study (Jolly & Reardon, 1985) reported that low analytics were slower in automating a simple skill. The limited evidence available suggests that low analytics might experience more difficulty during the strategy execution process. The difficulty might take the form of errors, requests for assistance (system help), and/or problems and uncertainty.

5.4.5 USER PERCEPTIONS

Implementation considerations will certainly require as much attention to user perceptions as to user performance. Most of the measures of user perceptions are obtained through the exit questionnaire. The confidence ratings are recorded in the user log during the system interaction.
5.4.5.1 CONFIDENCE

User confidence was measured with respect to both outcome (cost performance) and process. There is no basis for providing specific predictions on either item.

Despite my strong expectation that strategy-trained subjects will perform better, no difference in confidence has been predicted between the treatment groups because confidence ratings do not necessarily correspond to actual performance (Fischhoff, Slovic & Lichtenstein, 1977). In a training context, Cats-Baril & Huber (1987) found that the subjects who performed better reported lower confidence. For the strategy-trained subjects, confidence in process is also unavoidably, at least in part, a reflection of their confidence in me (the individual who presented the strategy). I have no basis for predicting what additional impact my involvement may have.

Research results concerning the relationship between CRA and confidence are also contradictory. DosSantos (1982) found that high analytics were more confident. Two other studies (Gul & Zaid, 1981; Gul, 1984) reported that low analytics were more confident.
5.4.5.2 PERCEIVED USEFULNESS OF INFORMATION

The exit questionnaire contains items concerning the perceived usefulness of information and items concerning specific problems related to the manner in which information was displayed. These items were included for exploratory purposes, and no specific predictions are made.

On the one hand, strategy-trained subjects should have all of the information they need to make their decisions. On the other hand, strategy-trained subjects should also be aware that a large part of the information provided by the system is not necessary for the solution of the problem. These perceptions contribute in opposite directions towards determining perceived usefulness, so I am unable to predict that strategy trained subjects would find the information more useful. Possible relationships with CRA are also unclear.

5.4.5.3 PERCEIVED DIFFICULTY

Perceived difficulty in separating relevant from irrelevant information is measured by an item on the exit questionnaire which simply asks how much
difficulty was experienced in determining what
information was important.

**Main effect of the training treatment:**

**H1(d):** Operations (O) trained subjects will report
greater perceived difficulty than strategy (S)
trained subjects:
(O > S).

For strategy-trained subjects, the relevant information
was identified as part of the training process. The
task of separating relevant from irrelevant information
was done for them, so the perceived difficulty should
be very low. In contrast, operations-trained subjects
were faced with a large number of information items to
evaluate.

**Interaction effect of training treatment with CRA:**

**H2(d):** High (H) analytic, strategy (S) trained
subjects will report the same perceived
difficulty as low (L) analytic, strategy (S)
trained subjects:
(HS = LS).

High (H) analytic, operations (O) trained
subjects will report lower perceived difficulty
than low (L) analytic, operations (O) trained
subjects:
(HO < LO).

High (H) analytic, strategy (S) trained
subjects will report lower difficulty than high
(H) analytic, operations (O) trained subjects:
(HS < HO).

For strategy-trained subjects, the relevant information
was identified as part of the training process. The
perceived difficulty of identifying the relevant
information should be very low for strategy-trained subjects and should not depend on cognitive restructuring ability because task demands for CRA were greatly reduced.

As discussed in Section 2.3.1, high analytics are better able to separate relevant from irrelevant information. Assuming that perceived difficulty corresponds to skill, the perceived difficulty on the part of high analytics should be less than for low analytics within the operations training treatment (Of course, the relationship between perceived difficulty and skill may be just as unpredictable as the relationship between confidence and performance.). Even though high analytics may be skilled in identifying relevant information, operations-trained high analytics faced an objectively more difficult task than strategy-trained high analytics.

5.4.5.4 OVERALL SATISFACTION WITH STRATEGY

Overall satisfaction with the strategy was measured by an item on the exit questionnaire which asked whether the subject would employ his/her strategy on the job if using the system were a job requirement. This item was included for exploratory purposes, and I have no basis for predictions of any kind.
5.4.5.5 PERCEIVED ADEQUACY OF DOMAIN KNOWLEDGE

While exposure to the required domain knowledge was partially controlled by providing duplicate lectures, subjects may well have differed in the extent to which they assimilated and retained that knowledge. Perceived adequacy of domain knowledge was measured by a questionnaire item which asked how much of the information presented by the system the subject felt he/she understood. This item was also included for exploratory purposes, and no specific predictions are made.

If a difference related to training treatment were to occur, I would expect that operations-trained subjects would perceive their domain knowledge as less adequate. For operations-trained subjects, the demands of formulating an appropriate strategy might make them comparatively more aware of inadequacies in their domain knowledge.

If a difference were to occur in the operations training treatment, I would expect high analytics to provide higher ratings of perceived adequacy than low analytics. High analytics may have been better attuned to information requiring an understanding of causal relationships (both conceptual and mathematical).
Additionally, low analysts who are expected to experience difficulty in formulating appropriate strategies might attribute that difficulty to inadequate knowledge.

5.4.6 MAINTENANCE AND GENERALIZATION OF TRAINING

The performance of subjects who participated in the verbal protocol sessions will be used to assess the maintenance and generalization of the suggested strategy. User logs and verbal protocols will be used to assess maintenance and generalization. For cost performance on the maintenance task, the same hypotheses presented above concerning cost performance outcomes will be applied (see Table 10).

For the generalization task, strategy-trained subjects are still expected to outperform operations-trained subjects ($H1(g): S > O$). However, given the difficulty of obtaining generalization, this prediction is an optimistic one.

For this task, differences related to CRA are expected to emerge even within the strategy training treatment, so a main effect of CRA is predicted ($H2(g): R > L$). Strategy training did not explicitly address generalization, although the subjects were provided with basic information concerning why and how the
suggested strategy worked. The demand for 
generalization requires that the subject engage in 
renewed strategy formulation, regardless of the 
training treatment. Hence, CRA will enter into 
performance for both treatment groups. These 
hypotheses, along with all of the other hypotheses 
stated in this chapter, are summarized in Table 10.
Table 6

**Subject Assignments to Treatment-by-Ability Groups**

<table>
<thead>
<tr>
<th>Treatment Condition</th>
<th>CRA</th>
<th>Row Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Strategy</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Operations</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Column Statistics</td>
<td>23</td>
<td>19</td>
</tr>
</tbody>
</table>

**Note.** High CRA subjects scored a 14 or higher on the GEFT (out of a possible 18). Low CRA subjects scored under 14.
Using the 7 point scale shown above, how sure are you that you found the lowest cost schedule possible for the job set you just completed? (Please respond with respect to schedules which can be created after attempting 100 or fewer changes.) 0

Using the same 7 point scale, how sure are you that the strategy you used was the best way to solve the problem if the objective is to create the lowest cost schedule possible while attempting 100 or fewer changes? 0

Figure 15. Confidence items screen.
Figure 16. Expected cost performance outcomes by treatment.
Table 7

Expected Differences in Information Item Use by Training Treatment

<table>
<thead>
<tr>
<th>Schedule Comparison Screen:</th>
<th>Job Sequences:</th>
<th>0 &gt; S</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subtotals by Cost Category:</td>
<td>0 &gt; S</td>
</tr>
<tr>
<td></td>
<td>Totals by Sequencing Rule:</td>
<td>0 &gt; S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Modification Screen:</th>
<th>Changeover Charges Column:</th>
<th>S &gt; 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tardiness Charges Column:</td>
<td>0 &gt; S</td>
</tr>
<tr>
<td></td>
<td>Inventory Charges Column:</td>
<td>0 &gt; S</td>
</tr>
<tr>
<td></td>
<td>Period Completed Column:</td>
<td>0 &gt; S</td>
</tr>
<tr>
<td></td>
<td>Subtotals by Cost Category:</td>
<td>0 &gt; S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analysis Screen:</th>
<th>Changeover Charges Column:</th>
<th>0 &gt; S</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inventory Costs per Period:</td>
<td>0 &gt; S</td>
</tr>
<tr>
<td></td>
<td>Periods Processing Required:</td>
<td>0 &gt; S</td>
</tr>
<tr>
<td></td>
<td>Tardiness Costs per Period:</td>
<td>0 &gt; S</td>
</tr>
<tr>
<td></td>
<td>Periods Late:</td>
<td>0 &gt; S</td>
</tr>
<tr>
<td></td>
<td>Net Effect:</td>
<td>S &gt; 0</td>
</tr>
<tr>
<td></td>
<td>Best Net Effect:</td>
<td>S &gt; 0</td>
</tr>
</tbody>
</table>

High/Low Values Screen: 0 > S

Changeover Costs Screen: 0 > S

Total Number of Information Items Used: 0 > S

Note. S = Strategy Trained. 0 = Operations Trained.
Table 8

Expected Differences in Information Item Use Within the Operations Training Treatment

<table>
<thead>
<tr>
<th>Modification Screen:</th>
<th>Subtotals by Cost Category:</th>
<th>H &gt; L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis Screen:</td>
<td>Net Effect:</td>
<td>H &gt; L</td>
</tr>
<tr>
<td></td>
<td>Best Net Effect:</td>
<td>H &gt; L</td>
</tr>
<tr>
<td>Changeover Costs Screen:</td>
<td></td>
<td>H &gt; L</td>
</tr>
<tr>
<td>Total Number of Information Items Used:</td>
<td>L &gt; H</td>
<td></td>
</tr>
</tbody>
</table>

Note. H = High CRA. L = Low CRA.
Table 9

Descriptions of High and Low Analytic Decision Processes

High Analytic Decision Process:

Behavior appears logical, methodical (structured) and purposeful. Jobs are moved following reasoning concerning the probable impact. Reasoning demonstrates the application of knowledge concerning causal relationships. Subject is able to predict system behavior, including net effects and total cost.

Low Analytic Decision Process:

Behavior appears to consist largely of trial-and-error. Jobs are frequently moved with no apparent justification. Subject neither applies knowledge of causal relationships, nor demonstrates the ability to predict system behavior.
Table 10 (continued on next page)

Summary of Predictions

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Performance</td>
<td>H1(p): S &gt; 0.</td>
</tr>
<tr>
<td></td>
<td>H2(p): LS = HS; HO &gt; LO; HS &gt;= HO.</td>
</tr>
<tr>
<td>Time Performance</td>
<td>No predictions made.</td>
</tr>
<tr>
<td>Information Use</td>
<td>H1(i): S &lt;&gt; 0. (see Table 7)</td>
</tr>
<tr>
<td></td>
<td>H2(i): LS = HS; LO &lt;&gt; HO. (see Table 8)</td>
</tr>
<tr>
<td>User Strategies</td>
<td>H1(s): S &lt;&gt; 0.</td>
</tr>
<tr>
<td></td>
<td>H2(s): LS = HS; LO &lt;&gt; HO.</td>
</tr>
<tr>
<td>Strategy Execution</td>
<td>H1(e): L &lt;&gt; H.</td>
</tr>
<tr>
<td>User Confidence</td>
<td>No predictions made.</td>
</tr>
</tbody>
</table>

Note. S = Strategy Trained.  O = Operations Trained.  H = High CRA.  L = Low CRA.
Table 10 (continued from previous page)

Summary of Predictions

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Usefulness of Information</td>
<td>No predictions made.</td>
</tr>
<tr>
<td>Perceived Difficulty of Determining Information Value</td>
<td>H1(d): O &gt; S.</td>
</tr>
<tr>
<td></td>
<td>H2(d): HS = LS; HO &lt; LO; HS &lt; HO.</td>
</tr>
<tr>
<td>Overall Satisfaction with Strategy</td>
<td>No predictions made.</td>
</tr>
<tr>
<td>Perceived Adequacy of Domain Knowledge</td>
<td>No predictions made.</td>
</tr>
<tr>
<td>Maintenance of Strategy Training</td>
<td>H1(m): S &gt; O.</td>
</tr>
<tr>
<td></td>
<td>H2(m): LS = HS; HO &gt; LO; HS &gt;= HO.</td>
</tr>
<tr>
<td>Generalization of Strategy Training</td>
<td>H1(g): S &gt; O.</td>
</tr>
<tr>
<td></td>
<td>H2(g): H &gt; L.</td>
</tr>
</tbody>
</table>

Note.  S = Strategy Trained.  O = Operations Trained.  \( H \) = High CRA.  \( L \) = Low CRA.
CHAPTER 6. DATA ANALYSIS

The first two sections of this chapter describe the experimental subjects and provide an overview of the analysis procedures, particularly the methods employed to conduct the protocol analysis. The remaining sections of this chapter present hypothesis tests and results in the order in which they were presented in the previous chapter.

6.1 DESCRIPTIVE STATISTICS

Of the 42 subjects included in the analysis, half were assigned to each treatment. Descriptive statistics for the sample are presented in Table 11. The average age was approximately 32 years, and more males than females took part (64% and 36%, respectively). These demographic characteristics were very similar for both treatment groups, although only the GEFT score was used in determining treatment assignments. The treatment groups had the same median GEFT score (14) and approximately the same mean GEFT score (12). The treatment groups were also very similar with respect to the distribution of GEFT scores, as shown in Figure 17.
Twelve subjects from the same course section provided concurrent verbal protocols. The 12 protocols were divided evenly between strategy-trained and operations-trained subjects. Descriptive statistics for the protocol group are provided in Table 12. Within each treatment group, two subjects were high analytics (GEFT scores of at least 14), and the remaining four were low analytics (GEFT scores of less than 14). The mean and median GEFT scores for the protocol group were lower than for the subjects as a whole. The protocol group contained the highest number of female subjects of any group, and female subjects scored lower on the GEFT ($p < .007$ for a two-tailed t-test). This sex difference has been reported in a number of different psychometric studies (Witkin et al., 1971; DeSanctis & Dunikoski, 1983). Table 13 summarizes the GEFT scores by sex.

6.2 OVERVIEW OF ANALYSIS PROCEDURES

This section provides background information concerning the analysis procedures employed. Each of the sections describing hypothesis tests provides additional information concerning the statistical tests employed.
6.2.1 STATISTICAL ANALYSIS PROCEDURES

A number of subjects were excluded from one or all of the analyses reported in this chapter. Appendix 4 lists all of the exclusions and the reasons for them. In addition, there was a natural attrition of subjects across data sets because subjects completed varying numbers of data sets in the time allotted for the experiment. The number of usable observations for each analysis was 42 or less (The number of observations will be reported for each analysis.).

With the exception of calculations concerning interrater reliability (which were completed using a spreadsheet package), statistical analyses were performed using the SAS software package for personal computers, Release 6.03 (SAS Institute Inc., 1988). Analysis of variance and regression analysis were performed using the General Linear Models procedure (proc GLM). SAS provides four types of hypothesis tests for the General Linear Models procedure. The significance levels reported here are for Type III hypothesis tests, which are appropriate for unbalanced designs (SAS Institute Inc., 1988, p. 588; Stevens, 1990, pp. 115-119). The Type III sum of squares represents the additional contribution of each variable
above and beyond the contributions of all other variables, which means the reported significance of each effect is adjusted for every other effect (SAS Institute Inc., 1988, p. 615).

In many cases, the data violated assumptions (e.g., normality, equal variances) underlying statistical tests. Therefore, nonparametric tests were used to confirm the pattern of results obtained through parametric tests. For example, both a t-test (proc TTEST) and a Wilcoxon rank-sum test (proc NPAR1WAY) were used to establish that the GEFT scores of males and females differed significantly. The significance levels reported in the text are for parametric tests unless otherwise indicated.

For analyses which required that subjects be classified as high or low analyitics, individuals scoring at the overall median of 14 or higher on the GEFT were classified as high analyitics. The median is often used as the dividing point in GEFT studies (Doktor & Hamilton, 1973; Lusk, 1973; Lusk, 1979; DosSantos, 1982; Gul, 1984). In addition, a GEFT score of 14 has been used as the dividing point in several other studies (Doktor & Hamilton, 1973; Benbasat & Dexter, 1979; Benbasat & Dexter, 1982; Benbasat & Dexter, 1985).
6.2.2 PROTOCOL ANALYSIS PROCEDURES

Transcripts of the verbal protocol sessions were generated from tape recordings. Full transcripts are presented in Appendix 5. The word repetitions and other idiosyncracies in the transcripts reflect the nature of the verbalizations rather than typographical errors. The transcripts are only identified by number, so an accompanying key indicates the subject’s treatment assignment, GEFT score and the delay between the experimental session and the individual protocol session. The transcript numbers merely indicate the order in which the tapes were transcribed.

Protocol analysis was undertaken at two levels of detail. The first level of analysis involved classifying protocols on the basis of full transcripts (aggregate analysis). The second level of analysis entailed coding the individual phrases within the transcripts (statement-by-statement analysis). Statistical tests were performed using the results of the statement-by-statement coding process. At both levels of analysis, two coders were used in order to determine the reliability of the coding procedures.

Two coders assisted me in the protocol analysis. Both coders were blind to the substance and purpose of
the study. The coders were trained at different times, but followed the same sequence of training and analysis procedures. The coder training/analysis procedures were as follows:

1) Experimenter provides training in job scheduling using the notes which guided the classroom sessions for experimental subjects.

2) Experimenter provides operations training.

3) Coder is given an opportunity to use the decision support system.

4) Experimenter provides strategy training.

5) Coder is asked to distinguish between the protocols of strategy-trained subjects and the protocols of operations-trained subjects (aggregate analysis).

6) Coder is asked to distinguish between the protocols of high analytics and the protocols of low analytics within the operations training treatment (aggregate analysis).

7) Experimenter provides training in the statement-by-statement coding scheme based on the protocols of strategy-trained subjects.

8) Coder is asked to assign phrases to coding categories for the protocols of operations-trained subjects (statement-by-statement analysis).

The first three stages of training were designed to provide the coders with an understanding of the task. Hayes (1982, p. 71) suggested that "it is wise to get
firsthand experience by doing the task yourself before trying to analyze protocols of other people."

The statement-by-statement coding scheme was developed to explore differences in the strategy formulation and strategy execution processes, particularly differences between high and low analytics. The first step in developing the coding scheme was to inspect the protocols to "extract the vocabulary of objects and relations needed to define the problem space and operators" (Ericsson & Simon, 1984, p. 264). I was guided by the descriptions of high and low analytic decision processes shown in Table 9 and the descriptions of the strategy formulation and strategy execution processes provided in Section 3.1. However, these general guidelines needed to be translated into the specific behaviors which occur in the context of the experimental task.

The initial coding scheme I developed failed to produce satisfactory interrater agreement. In addition, the number of protocol statements classified as miscellaneous ranged from 57% to 89% per protocol (based on the protocols of strategy-trained subjects used for training purposes). As a result of that first experience and the nature of the disagreements, I revised the coding scheme.
The revised (and final) coding scheme was developed using the protocols of strategy-trained subjects only. The first set of revisions included adding a number of categories for strategy execution activities in order to account more fully for protocol statements. No categories for strategy formulation activities were added. The coding definitions were made much more explicit, and a number of examples were added to the definitions.

The training procedure was also altered for the second coder in an attempt to improve the interrater agreement. The second coder wished to take a more active role during the training process. Rather than allowing me to illustrate the coding process before attempting it himself, he wished to attempt all statement classifications beginning with the very first training protocol. During training, the first coder had encountered each of the strategy-trained subjects' protocols only once. The second coder was trained to criterion (approximately 70% agreement on two consecutive protocols) using the strategy-trained subjects' protocols. While a more stringent criterion might have been set, 70% agreement was deemed sufficient in consideration of the very large number of coding categories involved (18 in all) and the length
of the protocols. Most of the protocols were coded twice before the criterion was achieved.

The coding procedure was also altered in order to increase interrater agreement. For the original coding attempt, both the coder and I had coded all of the transcripts before agreement was computed. The drawback of this approach was that consistent disagreements which indicated common errors or differences in interpretation of the coding definitions were in no way revealed or corrected. In order to prevent as many "avoidable" errors as possible, the revised coding procedure included a comparison of codings following each set of approximately 50 independently coded statements. In this fashion, the training procedure was extended beyond the initial set of protocols. The discussion of coding disagreements was also used to determine the final coding to be assigned for analysis purposes.

A list of the final coding categories, a description of the coding methodology, the coding definitions and a sample coding are provided in Appendix 6. These materials (with the exception of the sample coding) were also used during training. Because the coding process was based on semantic units, the way in which statements were numbered had little impact on
the coding process. A single numbered statement could include multiple coded phrases, or multiple numbered statements could be assigned a single coding. Analyses are based on the number of coding incidents rather than the number of statements unless otherwise indicated.

Each of the strategy formulation categories (1-10) can be mapped onto characteristic of a high or low analytic decision process (see Table 9). Categories 1 and 2 (method statement, specific move description) are considered evidence of a planned, structured decision process. Categories 3-5 (calculation, observation, information selection) are considered evidence of the activities of information selection, organization and reorganization. Categories 6-9 (prediction/knowledge of system behavior, explanation of system behavior, investigation/questioning of system behavior, domain knowledge statement/question) are considered evidence of predictive ability and/or attempts to base decisions on causal relationships. Category 10 (trial and error statement) is considered evidence of a low analytic decision process.

The strategy execution categories are concerned with processing feedback and managing user resources, especially memory. Category 11 (performance monitoring) is evidence of the processing of
performance feedback. Category 12 (progress monitoring) is evidence of the controlling of the execution process over time. Category 13 (interruption) suggests user uncertainty, difficulty and/or error. As discussed in Section 5.4.4, Category 13 is based on the simplifying assumption that all such incidents reflect failures during the strategy execution process. Category 14 statements (execution procedures) are evidence of the means users adopt to manage resources and/or control error rates. Categories 15-16 (operating help, strategy help) provide evidence concerning the means users adopt to respond to uncertainty/difficulty. Category 17 (forgetting) indicates that an attempt to retrieve information from memory has been unsuccessful.

In comparison to the initial coding process, the final coding process assigned between 16% and 36% of the protocol statements in any individual protocol to the miscellaneous category (based on the protocols of strategy-trained subjects used for training purposes). Overall, fewer than 12% of the total number of coding incidents were in the miscellaneous category (based on the final coding assigned for analysis purposes). The final coding scheme captured far more of the
information available from the protocols than the initial coding scheme.

Table 14 presents information concerning the reliability of the coding process based on the agreement between the second coder and me for the protocols of operations-trained subjects. For the 824 coding incidents, the overall percent agreement was 69%. For any individual protocol, the percent agreement ranged from 64% to 78%.

Percent agreement alone is not a good indicator of the reliability of the coding process because some agreement may occur by chance, so two other reliability indices have been provided. Cohen’s kappa is a common measure of interrater agreement which indicates the "proportion of agreement after chance agreement is removed from consideration" (Cohen, 1960, p. 40). The overall kappa coefficient of .65 indicates that approximately 65% of the coding incidents were agreements (with chance excluded). The Z scores are used to test the hypothesis that agreement was at chance levels (the population kappa = 0).

Perreault & Leigh (1989) suggest that the assumptions underlying Cohen’s kappa make it an "overly conservative measure of reliability, which in turn makes it difficult to compare values of kappa with
values from other types of reliability indices, such as Cronbach's alpha" (p. 138). They provide an alternative, less conservative measure of interrater reliability which is also presented in Table 14. The overall reliability index of .82 suggests that a typical judge could code about 82% of the total responses consistently.

In general, high agreement was difficult to achieve for categories which occurred infrequently. Statements from low frequency categories either did not occur during training or occurred so rarely that the coder had no basis for comparison when they occurred during the coding process. Low reliability for low frequency categories (i.e., Categories 3, 7, 8, 9, 10, 15, 17) is not of great concern, however, because these categories were not included in any statistical analyses.

The reliability for Category 18, miscellaneous, is very low because of the procedures used in counting disagreements and the types of problems which led to disagreement. Any time that a phrase was coded by one coder and not another, the error was assigned to Category 18. In some cases, the disagreement was "real" in the sense that the coder who did not code the phrase felt that the phrase did not match any of the
category definitions. Other real but less serious disagreements involved decisions concerning whether contiguous phrases contained multiple coding episodes or only one. Even when a series of phrases were assigned to the same coding category by both coders, disagreements concerning the number of separate episodes resulted in a large proportion of the Category 18 errors.

In addition to real disagreements, however, "avoidable" Category 18 errors occurred whenever lapses of attention occurred. A great deal of vigilance over extended periods of time is required to ensure that all of the relevant phrases have been coded. A large number of coding categories also have to be considered with respect to each phrase. In reviewing our coding assignments, one or the other of us frequently apologized for having missed something obvious. I would estimate that at least half of the Category 18 errors represented lapses in attention resulting from fatigue and/or boredom.

The real and avoidable Category 18 errors together constituted 69% of all coding errors. The second largest specific category of errors (excluding Category 18) was also related to lapses in attention. One of the subcategories within Category 14, rehearsal,
requires that the coder use preceding phrases to establish that information is being rehearsed rather than stated for the first time. Of the non-18 errors, 17% consisted of disagreements concerning rehearsal versus initial statement of information.

Some of the other coding disagreements resulted from ambiguities in the language of the protocols. For example, the referents of pronouns are not always clear, and the coders sometimes assumed different referents. Occasionally interpretations differed as to whether the letter to which a subject referred was a menu item or a job name.

The coding disagreements had to be resolved in order to perform subsequent analysis. Most of the disagreements were resolved through discussion (which was not difficult when one of us admitted a lapse of attention or memory). In cases where we agreed to disagree, I made the final decision as to the coding assignment. I had the advantages of extensive experience with the coding process, extensive experience with the system, and familiarity with subjects' behavior acquired through approximately 42 hours of observation.
6.3 COST PERFORMANCE

Tests of cost performance hypotheses are presented for the first three job sets only. Well under half of the subjects (36%) completed the fourth job set, and none of the strategy-trained, low analytic subjects completed the fourth job set. Only 26% of the subjects completed the fifth job set. The hypotheses were tested separately on each job set so that the maximum number of subjects could be included in each analysis. Cost performance scores for a job set were included only for subjects who had indicated that they were finished with that job set. The exception to this rule were three strategy-trained subjects whose performance for the first job set was recorded even though they had not indicated that they were finished when time expired.

In order to be able to compare performance across data sets, numeric cost performance was converted to a percentage as follows:

\[
\frac{\text{Starting Cost} - \text{Final Cost}}{\text{Starting Cost} - \text{Best Known Cost}}
\]

The resulting score indicates the percent of the maximum known improvement which was achieved. Figure 18 compares the average performance of the four
experimental groups across job sets. The relative ordering of average cost performance for the four experimental conditions was the same for all three job sets (from highest to lowest): LS, HS, HO, LO. Overall performance decreases and the differences between group means increase across job sets, as shown in Figure 18. Tables 15 through 17 present detailed performance data for the first three job sets. The GEFT data provided in the figures indicates that the decrease in overall performance occurred even though the average GEFT score remained stable across job sets (despite attrition).

For each data set, detailed analysis was undertaken only after an initial Kruskal-Wallis test and/or ANOVA suggested that at least two of the four groups differed in terms of the location of the distribution. For cost performance data, all three job sets passed the initial test. Table 18 presents the results of a regression analysis on the cost performance data for each job set. Figure 19 presents the graphs of the regression lines by treatment for each job set. The full model (treatment, GEFT score, and their interaction) explained between 25.6% and 42.9% of the total variance in cost performance data.

As anticipated, GEFT score alone did not contribute significantly to the prediction of cost
performance. There was a very pronounced impact of the training treatment on performance for every job set, whether parametric or nonparametric analyses were employed. As hypothesized (H1[p]), strategy-trained subjects outperformed operations-trained subjects on all three job sets. The GEFT data provided in Tables 15 through 17 suggest at most a one point advantage in mean GEFT score for the strategy-trained group (job set 3). However, the strategy-trained subjects were actually at a disadvantage with respect to GEFT when compared to their respective peer groups (high versus high and low versus low). Strategy-trained subjects outperformed operations-trained subjects of essentially the same ability level.

The expected interaction effect clearly reached statistical significance for the first job set (p < .04), but did not achieve accepted levels of statistical significance for the second and third job sets. However, the consistent pattern of results shown in Figure 19 provides evidence in favor of an interaction effect (Cronbach & Snow, 1977, p. 58). The direction of the coefficient for the interaction term is also consistent with the hypothesized direction for every job set. Especially because of the low power of this study (Cronbach & Snow [1977, p. 46] recommend
100 subjects per treatment in order to detect interaction effects.), the combined evidence can be taken as sufficient to reject the null hypothesis of no interaction.

The nature of the interaction was investigated in more detail by testing the specific contrasts making up interaction hypothesis H2(p) (see Table 10). The results of t-tests are presented in Table 19 (Wilcoxon rank sum tests suggested the same conclusions in 8 out of 9 cases.). As hypothesized (H2[p]), the results suggest an ordinal interaction in which high CRA leads to better performance only within the operations training condition.

The consistently higher performance of low analytics within the strategy training treatment was surprising to me. The difference approached significance for the first job set, and might well have obtained significance with a larger sample. In framing the cost performance hypotheses, I concentrated on the relationship between ability and performance. On the basis of users' ability to learn and utilize the suggested strategy, there was no reason to expect a performance advantage for low analytics.

Exit questionnaire data suggests, however, that the lower performance of high analytics was not the
result of an inability to learn and utilize the suggested strategy. Rather, three high analytic subjects indicated that they had intentionally altered the strategy to bring about improvements. These were the only subjects to indicate that substantive changes in the strategy had been made, and these subjects did not perform as well with their modified strategies as they would have if they had completely followed the suggested strategy. Thus, the performance difference appears to have been related to subjects' willingness to use the suggested strategy rather than their ability to do so.

6.4 TIME PERFORMANCE

As with cost performance hypotheses, tests of time performance hypotheses were performed for the first three job sets only. Time performance data was included for a job set only for subjects who had indicated that they were finished with that job set. Again, the only exceptions were the three strategy trained subjects who did not complete the first job set. All three were assigned a time performance score of 3600 seconds.

The group time performance means across job sets are graphed in Figure 20. The design of the study
makes interpretation of the results for all but the first job set problematic. Because subjects were given a fixed time period in which to work, subjects who required the longest times per schedule dropped out of later data sets. As a result, each job set reflects the time performance of a faster and faster group of subjects. The convergence across time exhibited in Figure 20 may be simply a byproduct of subject attrition.

For all three data sets, the rank ordering of the means (from highest to lowest time required) began with LS, followed by HS subjects. However, an initial Kruskal-Wallis test and ANOVA suggested that the groups differed significantly only for the first job set. Table 20 summarizes the average time performance for each group on the first job set.

A regression model with GEFT as a continuous variable did not predict a significant proportion of the variance in time performance on the first job set. However, an analysis of variance suggested main effects of both training treatment and CRA, as shown in Table 21. Wilcoxon rank sum tests also yielded significant results for both main effects (p < .06 for treatment, and p < .04 for CRA).
The results of the analysis of variance suggested that strategy-trained subjects required more time to complete the first job set. An additional test was run, however, in order to determine whether the results for the first job set were attributable to the three outlying strategy-trained subjects who were assigned times of 3600 seconds. Two of the subjects were low analytics (GEFT scores of 3 and 12). The next highest time performance for a LS subject was 3147 seconds (approximately 7.6 minutes less). One of the subjects was a high analytic (GEFT score of 14). The next highest HS subject required 2434 seconds (approximately 19.4 minutes less). Without the three extreme subjects in the analysis of variance, the difference between the treatment groups was no longer significant for the first job set.

For the first job set, low analytics required more time on the average than high analytics (as suggested by a value of p < .06 for the analysis of variance). The effect of CRA was significant even when the subjects with scores of 3600 were omitted from the analysis, although at a reduced significance level (p < .10). The pattern of results is consistent with the idea that low analytics experience particular difficulty during the early, strategy formulation
process but less difficulty thereafter. Of course, other plausible interpretations might be proposed.

In summary, the data provided limited evidence of time performance differences between groups for the first data set only. It is impossible to determine if the differences in time performance found for the first job set would have persisted over time if all of the subjects had been allowed as much time as they required to complete all of the job sets.

In order to test whether the amount of time spent in making decisions resulted in higher performance, a regression analysis was conducted in which treatment group, the interaction between treatment and GEFT score, and time performance were used to predict cost performance. Once the other effects in the model had been taken into account, time contributed marginally to performance (p < .08) for the second job set only. Further analysis suggested no evidence of an interaction effect between treatment or GEFT score and decision time. Thus, the relationship between decision time and performance was weak to nonexistent.
6.5 THE STRATEGY FORMULATION PROCESS

The analyses reported in this section are intended to provide insight into the process differences which accounted for the performance differences discussed in Section 6.3.

6.5.1 INFORMATION USE

The information use patterns were determined from the usage items on the exit questionnaire. Except for the page requesting information on the High/Low Values Screen, when an item was not marked, the value was recorded as missing. For the High/Low Values Screen and the Changeover Costs Screen pages of the exit questionnaire, a blank page was interpreted as indicating that the screen had not been used. All subjects were included in this analysis.

In all cases, the items were coded as marked on the questionnaire, although the written comments in a few cases contradicted the marked responses or suggested that the question may not have been interpreted as expected. There appears to have been some confusion about the word "use" on the exit questionnaire. I intended the word "use" to signify that the item had entered into the decision in the
sense of influencing the actions to be taken. Written questionnaire responses indicated that a more liberal interpretation of "use" may have been employed by some subjects. For example, items may have been marked as used simply because their values had been noted. Three strategy-trained subjects indicated that they had used items on the Schedule Comparison Screen while simultaneously writing "I really did not use the data" when asked to explain how the information had been used. Another interpretation of "use" which appears in written comments was to indicate that information had been used even when it was used on a different screen from the one in question. For example, one strategy-trained subject indicated that the job sequence information on the Schedule Comparison Screen was used "when eliminating the negatives," actions which take place at the Analysis Screen.

The power of the statistical tests used in this section (chi square and Fisher's exact test) depend on both the relative magnitudes of the expected proportions and the number of subjects. Even for some of the tests using the full sample, the likely power of the test is below 70%. The problem is very severe with respect to associations within the treatment groups between CRA and information use. Even though there
is a danger of finding spurious results given the large number of tests which were performed, a significance level of .10 was deemed appropriate given the problem of low power and the exploratory nature of the analysis.

6.5.1.1 Expected Differences by Training Treatment

The detailed predictions concerning information item use by treatment were presented in Table 7. Table 22 summarizes the corresponding results with respect to the association between item use and treatment. For the Schedule Comparison Screen, the only significant difference between groups in item usage was in the direction predicted: operations-trained subjects were more likely to use the subtotal information than strategy-trained subjects. Overall, strategy-trained subjects reported using an average of 1.3 items (out of a maximum of 3), while perfect adherence to the suggested strategy would result in no usage at all.

For the Modification Screen, 3 of the 5 hypothesized treatment differences were statistically significant (see Table 22). Although the direction was correct, no significant difference was found for the changeover charges column, where usage was high regardless of treatment (57% of operations-trained
subjects, 80% of strategy-trained subjects). Also, no difference was found for the subtotals, where usage was in the expected direction but low regardless of treatment (33% for operations-trained subjects, 25% for strategy-trained subjects). The average number of items used by strategy-trained subjects was 3.1 (compared to the suggested strategy value of 3), while operations-trained subjects used an average of 5.1 items (out of a maximum of 7).

For the Analysis Screen, 5 of the 7 hypothesized differences were found. Reported usage of net effect and best net effect were high regardless of training treatment (in terms of the proportion of subjects reporting using the information at all). Strategy-trained subjects used an average of 2.0 information items (in keeping with the suggested strategy), and operations-trained subjects used an average of 4.2 items (out of 7).

Relatively few people used either the High/Low Values Screen or the Changeover Costs Screen. As predicted, operations-trained subjects used more items from the High/Low Values Screen. In fact, in keeping with the suggested strategy, none of the strategy-trained subjects used this screen at all. Significant associations between treatment and individual High/Low
item use were found for the four out of seven items where use by operations-trained subjects was high enough (not shown individually in Table 22). No significant difference was found for the Changeover Costs Screen, where use was low regardless of treatment.

As predicted, operations-trained subjects used significantly more information items from the system as a whole than did strategy-trained subjects (an average of 13.2 items for operations-trained subjects versus 6.6 for strategy-trained subjects). Average information use by strategy-trained subjects was slightly higher than the suggested value of 4 items.

6.5.1.2 Information Use within Training Treatments

Five predicted differences between low and high analytics were predicted within the operations training treatment (see Table 8). No significant differences between ability groups were expected in the strategy training treatment. All information items were tested for differences, and the few differences which were found are reported in Table 23.

Only two significant differences in item usage were detected within the operations training treatment, both of which were predicted. Operations-trained high
analytics were more likely to report using the net effect information and the Changeover Costs Screen. The direction of the difference was as expected for the subtotal information on the Modification Screen, but the difference was not significant (45% use for HO, 20% use for LO).

For the strategy-trained subjects, the only significant differences in item use involved the Schedule Comparison Screen. Low analytic strategy-trained subjects were more likely than HS subjects to use information from this screen. Low analytic subjects reported using a greater number of items from the Schedule Comparison Screen, but high analytic subjects reported using a higher number of items from the Modification Screen. There were no significant differences in the total number of items used from the entire system.

In summary, strategy training had a significant impact on information use. Hypotheses based on the nature of the suggested strategy were usually confirmed. There were no significant differences which had not been predicted. Alternative hypothesis H1[i] was supported. The average number of items used by strategy-trained subjects closely resembled the number required by the suggested strategy, which indicates
that the suggested strategy was followed in terms of the information considered. Strategy-trained subjects performed better than operations-trained subjects while using fewer information items.

The impact of CRA on information use was only partially in keeping with hypothesis H2[i]. Within the operations training treatment, it does not appear that high analytics were more selective in their information use. For all screens, high analytic, operations-trained subjects reported using a higher average number of items than LO subjects, although the difference was only significant for the Schedule Comparison Screen. This result is not consistent with a view of high analytics as more selective users of information. However, it is consistent with the finding that high analytics sample more fully from the available cues (Goodenough, 1976).

The low power of the statistical tests performed within treatment groups may have masked some differences related to CRA. On the other hand, there were some unexpected differences within the strategy-trained group, so there is limited evidence of a difference in information use related to CRA within the strategy training treatment.
6.5.2 USER STRATEGIES

This section will consider data from two main sources: concurrent verbal protocols and the user logs. Questionnaire data will be used to support some of the results.

6.5.2.1 CONCURRENT VERBAL PROTOCOLS

The purpose of the first aggregate protocol analysis was to verify that the experimental treatments resulted in recognizably different GOMS as evidenced by user behavior. Both raters were asked to distinguish between the protocols of strategy-trained subjects and the protocols of operations-trained subjects. The raters were not told how many subjects to expect in each treatment.

I censored the transcripts of strategy-trained subjects prior to the analysis to remove references to strategy help. I also attempted to remove or modify blatant references to strategy training or "stages." The modified transcripts of strategy-trained subjects appear in Appendix 7. The transcripts of operations-trained subjects were not altered in any way.

I sometimes experienced difficulty in determining the extent to which statements should be modified. For
example, does terminology such as "I have to..." or "I can only..." necessarily imply that strategy training was received? I decided that it does not. Overall, I attempted to modify transcripts as little as possible. The comments of the first rater indicated that subtle language cues were used in classifying one transcript. The comments of the second rater focused exclusively on actions the subjects performed. The rater classifications appear in Table 24.

Of the 12 classifications attempted, 75% were correct for both raters. Using Perreault & Leigh's (1989) index of reliability with the actual treatment assignment as one of the "judges", the reliability with which the treatment could be identified was .88 (Cohen's kappa of .67). These findings provide support for alternative hypothesis H1[s]. There were identifiable differences in the strategies used by the different treatment groups.

The treatment misclassifications are also of interest. Two of the three errors involved classifying operations-trained, high analytics as strategy-trained subjects (Transcripts 2 and 5 in Appendix 5). In both cases, the subjects had independently devised a method for using the net effects which resembled the suggested strategy. The other error was to classify a
low analytic, strategy-trained subject as an operations-trained subject (Modified Transcript 6). The subject exhibited difficulty in recalling the suggested strategy. This subject had not attended the lecture on job scheduling (which she also mentioned herself on the exit questionnaire). An inadequate knowledge base may have added to the difficulty of meaningfully encoding the strategy in the first place.

Although the detailed coding scheme was developed primarily to explain differences between high and low analytics, the results of a statement-by-statement analysis were also used to investigate differences in strategy formulation activity between treatments. After each statement had been classified according to the coding scheme in Appendix 6, subject profiles were developed which indicated the percentage of all coding episodes for each subject in each category. This method was used to enable comparison of protocols of varying lengths (from 17 to 242 numbered statements).

As discussed in Section 6.2.2, the first 10 coding categories are most relevant to the strategy formulation process. Categories 3, 7, 8, 9 and 10 were not analyzed individually because of very low rates of occurrence (average use less than 1%). The other
categories (1, 2, 4, 5 and 6) were analyzed individually using two-tailed Wilcoxon rank sum tests. No significant differences were found between treatment groups when individual categories were analyzed. However, when the percentage of verbalizations devoted to each of the first 9 categories was summed, operations trained subjects exhibited significantly higher rates of strategy formulation activity (34% of coding incidents for strategy-trained subjects, 54% of coding incidents for operations-trained subjects; p < .04). Because strategy formulation demands were reduced for strategy-trained subjects, it is appropriate that they should devote relatively less of their verbalization to strategy formulation activities.

The results of the statement-by-statement analysis were also used to explore differences in strategy formulation activity between high and low analytics. For the coding categories with rates of occurrence above 1% (1, 2, 4, 5 and 6), one-tailed Wilcoxon rank sum tests were used to investigate the relationship between CRA and percent of coding incidents in those categories. It was hypothesized that high analytics would exhibit more "analytic" behavior. Because each of the first nine coding categories was designed to capture dimensions of analytic behavior (see Section
6.2.2), it was hypothesized that high analytics would provide more verbalizations in each of the categories analyzed. While it would have been desirable to investigate the relationship of CRA to behavior within each treatment group, the extremely small number of subjects prohibited any analysis beyond main effects.

The results of the statement-by-statement analysis with respect to CRA are presented in Table 25. High analytics did have higher relative percentages of coding incidents in four of the five categories analyzed. The differences were significant for two categories. High analytics provided significantly more method statements (Category 1), which suggests that they employed a more structured decision process. High analytics also made more observations (Category 4), which suggests that they more actively organized the information available from the system. Overall, there was a significant difference between high and low analytics in the percent of coding incidents devoted to strategy formulation activities (the sum of the first nine coding categories). For high analytics, an average of 54% of all coding incidents were devoted to strategy formulation activities, as compared to 38% for low analytics. The protocols of high analytics did suggest a more analytic decision process.
Statement-by-statement analysis results were not used to test interaction effects because of the extremely low number of subjects in each treatment (six). However, the second aggregate analysis was designed to detect strategy differences between high and low analytics. The raters were given the protocols of operations-trained subjects and asked to classify them as high or low analytics based on the process descriptions in Table 9.

Of the 6 classifications attempted, 5 were correct for both raters (83% correct classifications). Using Perreault & Leigh's (1989) index of reliability with the true level of CRA as one of the "judges," a subject's analytic disposition could be identified with 92% reliability. The corresponding value of Cohen's kappa was .72.

The protocols provided additional evidence of a difference between the strategies of operations-trained high and low analytics: the manner in which net effect information was incorporated into the decision. Table 26 summarizes the methods subjects used to respond to net effect information. As indicated in the discussion of the first aggregate analysis, the two high analytic subjects used the net effect in much the same fashion which had been suggested during strategy training. The
low analytics, on the other hand, employed a variety of methods for responding to the net effect information which had little relationship to its intended use. Subjects were willing to develop methods for using net effect information even when they admitted that they did not understand what the number signified.

The first page of the questionnaire provides insight into the nature of the strategies used by all of the operations-trained subjects. In order to determine how important the net effect information was to the overall strategy, each strategy description was examined for mention of the net effect information. The assumption was that if the information were central to the decision process, it would be explicitly mentioned. A Fisher’s exact test was performed on the resulting frequency counts. It showed that for all operations-trained subjects, high analytics placed greater importance on the net effect information than low analytic subjects (p < .03). This replicates the result found for use of the net effect information item, and reinforces the differences observed in the protocol subgroup.

Three out of four of the low analytic, operations-trained subjects (transcripts 8, 10 and 12) exhibited a behavior which neither of the high analytic subjects
exhibited. These three subjects would try a long series of moves without reversing the moves which had resulted in large increases in cost. They apparently did not view their cost performance as dependent on each individual action taken. In comparison, the high analytics immediately reversed unsuccessful changes.

In summary, there was evidence that high and low analytic subjects used different strategies within the operations training treatment. The other aspect of the interaction suggested in H2[s] is the equivalence of the strategies employed by high and low analytic subjects in the strategy training treatment. One measure of the uniformity of strategies employed by strategy-trained subjects is the extent to which the subjects themselves perceived that they had followed the suggested strategy (the first item on the exit questionnaire for strategy-trained subjects). On a seven point scale (seven indicating total compliance), high analytics provided an average rating of 5.8. Low analytics provided an average rating of 5.1. The difference was not significant. The high and low analytics did not perceive themselves as behaving differently. The computer usage logs provide a more detailed means of comparing behavior.
6.5.2.2 USER LOGS

Three types of behaviors related to user strategies were analyzed: frequency of net effect operator use, frequency of best net effect operator use, and the number of calculations performed in the course of making a decision. The number of calculations performed is equivalent to the number of alternatives explored when moves are not repeated. In practice, my observations indicate that moves were frequently repeated.

The hypotheses were tested separately on each job set so that the maximum number of subjects could be included in the analysis. Operator counts for a job set were included only for subjects who had indicated that they were finished with that job set (for the first job set as well). For each operator type, detailed analysis was undertaken only after an initial Kruskal-Wallis test and/or ANOVA suggested that at least two of the four groups differed in terms of the location of the distribution.

6.5.2.2.1 NET EFFECT OPERATOR USE

The initial Kruskal-Wallis and ANOVA tests were significant for all three job sets. The results of
a regression analysis are presented in Table 27. For all job sets, strategy-trained subjects used the net effect operator more intensively than operations-trained subjects. Analysis of the computer log data provides evidence of a strategy difference which was not apparent through analysis of the questionnaire usage data alone (no difference in the proportion of users reporting usage was detected). For all job sets, the average number of times strategy-trained subjects invoked the net effect operator (25.9, 23.7 and 31.0) was very close to the number of times I invoked it when completing the same job sets (24, 30 and 36). Again, the behavior of strategy-trained subjects is in accordance with the suggested strategy.

Unlike the cost performance data, the net effect operator suggests a main effect of CRA. High analytics used the operator more frequently regardless of treatment. The interaction effects which were somewhat elusive for the cost performance data are quite pronounced with respect to net effect operator use. Table 28 summarizes the results of Wilcoxon rank sum tests. The differences between high and low analytics in the operations training treatment were significant in the expected direction for every job set. On the contrary, there were no significant differences in the
frequency of operator use between high and low analytics in the strategy training treatment. This confirms the nature of the interaction set forth in H2[s].

6.5.2.2.2 BEST NET EFFECT OPERATOR USE

The initial Kruskal-Wallis tests were highly significant for all three job sets. The results of Wilcoxon rank sum tests on specific contrasts are presented in Table 29. Beyond the first data set, the results mirror those found for the net effect operator. Strategy-trained subjects used the best net effect operator more often, high analytics used the best net effect operator more often, and the difference between high and low analytics was significant for the operations-trained subjects only. Again, the average frequency with which the operator was invoked by strategy-trained subjects (2.6, 2.5 and 3.1) was very close to the number of times I invoked it (3, 3 and 4).

6.5.2.2.3 CALCULATIONS

The initial Kruskal-Wallis and ANOVA tests were significant for the first two job sets. The results of a regression analyses are presented in Table 30. No predictions concerning the number of calculations were
made. The treatment had a significant effect on the number of calculations performed, as confirmed by Wilcoxon rank sum tests (p < .005 for the first job set, and p < .0003 for the second). Strategy-trained subjects performed more calculations than operations trained subjects for the first two job sets. The average numbers of calculations performed by strategy-trained subjects (34.6, 25.7 and 20.6) were close to the numbers I performed (31, 34 and 24).

For the first two job sets, the higher performance of strategy-trained subjects may have been due to a more thorough consideration of alternatives. For the third job set, however, the number of alternatives considered did not differentiate between treatment groups even though performance was different. In fact, for operations-trained subjects there were no significant correlations (p < .10) between the number of calculations performed and cost performance. The lack of correlations may suggest that ineffective search processes were being employed. In comparison, the number of calculations performed by strategy trained subjects did correlate with cost performance (p < .03 for job set 1; p < .02 for job set 3).

In summary, The protocol and user log data clearly suggest differences between the strategies of
strategy-trained and operations-trained subjects (H1[s]). The same data sources also suggest differences between the strategies of low and high analytics. The nature of the proposed interaction effect is supported (H2[s]). Considerable differences between the strategies of low and high analytics were detected within the operations training treatment. There was no evidence of systematic differences in the behaviors of low and high analytics within the strategy training treatment.

6.6 THE STRATEGY EXECUTION PROCESS

The same data sources used to explore strategy formulation differences also provide information concerning potential strategy execution differences.

6.6.1 CONCURRENT VERBAL PROTOCOLS

Coding categories 11-17 were designed to reflect differences in the nature of the strategy execution process. The last three categories were not analyzed because of average rates of occurrence under 1%. Two-tailed Wilcoxon rank sum tests on the remaining categories failed to provide evidence of treatment differences in the strategy execution process.
The only specific prediction regarding a main effect of CRA on strategy execution was that low analytics would experience more uncertainty/difficulty than high analytics. Uncertainty/difficulty was coded as Category 13 (interruption). A significant effect (at approximately $p = .10$) was found for a one-tailed test of Category 13. Low analytics interrupted the course of processing more often in order to solve problems or express uncertainty (14% of coding episodes for low analytics, 8% of coding episodes for high analytics).

For the remaining categories, two-tailed tests were employed. Only one significant ($p = .10$) difference emerged: high analytics spent more time monitoring their progress (Category 12) than low analytics (8% of coding incidents for high analytics, 3% for low analytics). My impression is that high analytics were more careful in avoiding duplicate moves and spent accordingly more time in keeping track of what had already been done.

6.6.2 USER LOGS

Error data and help usage were analyzed as possible indicators of difficulty/uncertainty. The error rates and help usage rates were generally low.
Kruskal-Wallis tests were used to determine whether to conduct further analyses for error and operating help data. The only indication of differences was for error rates on the first job set.

For the first job set, one-tailed Wilcoxon rank sum tests indicated that low analytics committed significantly more errors than high analytics, both overall (p < .03) and within the strategy training treatment (p < .005). Low analytics also committed more errors in the operations training treatment, but the difference was not significant. No other significant differences in error rates or help usage rates were found.

In summary, there was no evidence of a treatment effect on the strategy execution process. There was support for alternative hypothesis H1[a]. In particular, both the protocol analysis results for Category 13 and the error rate data suggest that low analytics experienced greater difficulty/uncertainty than high analytics. There was no indication of an interaction effect.

6.7 USER PERCEPTIONS

User perceptions were measured in order to determine if there were any negative affective
responses to training. Most of the data was collected through the exit questionnaire.

6.7.1 CONFIDENCE

For each of the first three job sets, confidence in the decision outcome was slightly (but not significantly) higher than confidence in the decision process used. Kruskal-Wallis tests were used to determine if there were any differences in confidence ratings between the four experimental groups. The only confidence items which passed the initial significance test \( p < .10 \) were the items for confidence in decision outcome for the first two job sets. Subsequent contrasts were analyzed using Wilcoxon rank sum tests. The tests were two-tailed since no predictions were made concerning the confidence results.

For every confidence item, the two highest values for average confidence were awarded by the strategy-trained low and high analytics. For the first job set, there was a main effect of the training treatment \( p < .003 \) on confidence in the decision outcome. For the second job set, the only significant contrast was between the strategy-trained and operations-trained low analytics \( p < .02 \). The results suggest that strategy
training led to higher confidence, particularly for low analytics.

6.7.2 PERCEIVED USEFULNESS OF INFORMATION

The items concerning perceived usefulness were standardized (mean 0, standard deviation 1) and then combined into indices of perceived importance and perceived usableness as described by Larcker & Lessig (1980). No significant differences were found between groups on either index based on Kruskal-Wallis tests. The items were also analyzed individually. Initial Kruskal-Wallis tests suggested that three of the five items differed across the experimental groups.

Subsequent tests of contrasts revealed differences related to CRA, but no differences related to the training treatment.

6.7.2.1 PERCEIVED IMPORTANCE

Wilcoxon rank sum tests indicated significant differences in perception concerning the sufficiency of the data presented (item C on page 7). High analytics expressed higher levels of agreement that the data was sufficient, both overall (p < .009) and within the operations training treatment (p < .03). However, elsewhere on the exit questionnaire, high analytics
made far more suggestions as to information or capabilities which should be added to the system (15 suggestions from high analytics and only 1 from a low analytic). The two results can be reconciled in that the majority of suggestions were for "convenience" features rather than new information. Apparently the high analytics felt that the information presented by the system was sufficient, but not necessarily in the most convenient form for their purposes.

High and low analytics also differed in their perceptions of how much of the data presented was essential to the decision (item B, page 8). Low analytics felt that a greater proportion of the data was essential. The difference was significant overall \( p < .03 \) and within the strategy training treatment \( p < .02 \). While the information use data suggested that high analytics considered more information items than low analytics, this questionnaire item indicated that they considered fewer of the items to be essential. Hence, while high analytics may have sampled more fully from the available cues, they may also have been better able to isolate those few cues which were most relevant (essential).
6.7.2.2 PERCEIVED USABILENESS

Wilcoxon rank sum tests of contrasts indicated that the perceived complexity of mental calculations required to use the data differed by CRA within the strategy training treatment. Low analytics indicated greater perceived complexity than high analytics (p < .03) in response to item B on page 7 of the questionnaire. Low analytics provided an average rating of 3.3 on a seven point scale with 7 denoting high perceived complexity. The average rating for high analytics was 1.8.

6.7.3 PERCEIVED DIFFICULTY

Low and high analytics differed in their perceptions of how much of the data presented by the system was essential to the decision. Despite these differences, a Kruskal-Wallis test did not suggest differences between any of the experimental groups in the perceived difficulty of separating relevant from irrelevant information (item D, page 7). Both high and low analytics reported problems related to the appropriate use of information (seven problems in all). Even within the strategy training treatment, a few
problems of information use were reported (three in all).

For the task as a whole, strategy-trained subjects were asked to describe specific areas of difficulty and/or confusion on the first page of the exit questionnaire. Half of the high analytics indicated specific areas of difficulty and/or confusion, and one third of the low analytics reported specific areas of difficulty and/or confusion. While strategy-trained, low analytics reported greater perceived complexity of the calculations required to use the suggested strategy, this was not reflected in cost performance or the responses concerning difficulty.

6.7.4 OVERALL SATISFACTION WITH STRATEGY

The initial Kruskal-Wallis test for differences between groups in overall satisfaction (item E on page 7) was highly significant (p < .003). Subsequent Wilcoxon rank sum tests of specific contrasts suggested both main effects and a possible interaction. Table 31 presents the average ratings provided by each group.

Strategy trained subjects were significantly more satisfied with their strategies than operations trained subjects (p < .02). High analytics were significantly more satisfied with their strategies than low analytics
(p < .02). No significant difference was found between high and low analytics in the strategy training treatment. However, high analytics were more satisfied with their strategies than low analytics in the operations training treatment (p < .003).

The satisfaction item also provides an indication of the extent to which subjects took the experiment seriously. Although it is only a hypothetical question, the realism of the experiment is tested by asking the subjects if they would apply their strategies on the job. The overall average rating of 4.4 (with a maximum of 7) suggests that the subjects were satisfied enough with their strategies that they would consider applying them on the job.

The verbal protocols provide further evidence that subjects attended to the experimental task. During the second protocol task, one subject said, "I am not going in to see my boss with the finished schedule right now or I won't be working on this any longer" (Transcript Extension 3, line 143). The same subject (Transcript 3, line 64) referred to the motivation provided by extra credit. Even days after the experimental session, three subjects (Transcripts 1, 3 and 5) recognized a job set they had encountered before, suggesting that they had attended to the task.
6.7.5 PERCEIVED ADEQUACY OF DOMAIN KNOWLEDGE

A Kruskal-Wallis test suggested no significant differences between the experimental groups in perceptions of the adequacy of domain knowledge. The overall average rating was 5.0 on a seven point scale where seven indicates that all of the information items were understood. The subjects had opportunities during the class lecture and training sessions to ask questions concerning the domain, but apparently emerged from both sessions with an imperfect understanding. While the lack of total understanding is lamentable, there is no reason to believe that imperfect understanding is unique to this experimental situation, or unique to experimental situations in general.

While all of the subjects were MBA students, the lack of understanding and/or application of domain knowledge was at times surprising. For example, all else constant, a high inventory holding cost per period suggests that a job should be moved towards the beginning of the sequence. Tardiness costs also imply directional moves. However, subjects sometimes moved jobs in the wrong direction in reaction to the information or moved the job randomly in either direction. Some subjects failed to discriminate
between constant (per period) and variable information (computed charges). Very few subjects reported using subtotal information, one of the few means of understanding cause and effect relationships.

Equally as worrisome as the fact that subjects did not understand basic information was their willingness to use it anyway. As indicated in Table 26, subjects based decisions on information which they admitted that they did not understand. Even faulty reasoning can lead to occasional positive results, particularly for this task. The reinforcement provided by cost improvements, however infrequent, seemed to be sufficient for poor strategies to be maintained. As noted by Brown et al. (1983, p. 91), inferior strategies resulting in partial success can impede development of better strategies.

6.8 MAINTENANCE & GENERALIZATION OF TRAINING

Section 6.3. presented evidence concerning the superior performance of strategy-trained subjects. This section uses data collected during the protocol sessions to determine whether that superiority persisted over time and/or transferred to a variant of the experimental task. Because of the small number of subjects, only tests of main effects were performed.
6.8.1 MAINTENANCE

The individual protocol sessions were conducted from two to nine days after the original training session. The first job set for the protocol session corresponded to the third job set for the experimental session. Using difference scores for the change in performance over time, average performance decreased almost significantly ($p < .101$) for the group as a whole (based on a two-tailed, parametric test for a mean difference of 0). The change in performance scores over time was not significantly related to CRA, although the magnitude of the decrease was greater for low analytic subjects. There was no significant difference between training treatments in the extent to which performance decreased between sessions, although the magnitude of the decrease was twice as large for strategy-trained subjects.

The behavior of strategy-trained subjects during the delayed sessions was sufficiently influenced by the suggested strategy that they could be identified through aggregate analysis of their protocol transcripts as described above. Strategy trained subjects still performed better than operations trained subjects during the delayed session ($p < .05$ for a two-
tailed, Wilcoxon rank sum test). During the experimental session, the strategy-trained protocol subjects outperformed their operations-trained counterparts by approximately 41%. During the delayed session, the difference in average performance dropped to approximately 31%. The user logs indicate that strategy-trained subjects made more frequent use of the net effect and best net effect operators, both during the immediate and delayed sessions. There were no significant differences in time performance or the number of calculations performed, either for the immediate or delayed sessions.

I can say without hesitation that the strategy-trained subjects attempted to use the suggested strategy. The notes I created during the protocol sessions indicate a variety of deviations from the suggested strategy which accounted for the degradation in performance. The deviations were more pronounced for the second and third stages of the strategy. These stages were also mentioned in written comments on the exit questionnaire as particularly difficult or confusing. The stages can be reconstructed if the subject remembers what the method is supposed to accomplish, but only one subject (Transcript 3)
exhibited a good understanding of how and why the strategy worked.

I believe that the best way in which to improve the maintenance of the suggested strategy over time would be to allow for more extensive practice and individual attention. Because practice time was severely limited, I did not have as much time as I needed to respond to questions and verify that the individual had understood. Subjects had difficulty in remembering the procedures which were difficult or confusing at the time of training. Overlearning has improved subsequent retention in other training contexts (Baldwin & Ford, 1988).

6.8.2 GENERALIZATION

As described in the previous chapter, the second sequencing task included a constraint which the subjects had not previously encountered: job A was not allowed to be late. I intentionally selected a constraint which would be violated if the suggested strategy were followed without modification. The protocol transcripts for the generalization task are presented in Appendix 8.

I did not anticipate that subjects would have difficulty in understanding the constraint, but at
least some of them did (Transcripts 8, 9 and 10). The difficulty was primarily caused by a lack of understanding of what "lateness" meant in terms of the information presented by the system and how it was affected by user actions. Both lateness and tardiness charge information could be used to determine if the constraint was met, but most subjects used only one or the other. One subject even used a system error message to determine if the constraint had been met: she repeatedly indicated that she was finished with the schedule in order to determine the status of job A. The system had been designed to prevent termination of a job set when the constraint was not met.

My observations indicated that at least five of the subjects experienced difficulty related to the use of the "restore best" operator. Because the lowest cost schedule was not necessarily a feasible one, subjects could "lose" a feasible solution by restoring the lowest cost schedule. The performance scores were based on the feasible solutions which subjects had attained at the time that they indicated that they were finished. At least two of the subjects "lost" feasible solutions which were better than their final solutions.
I modified my strategy for approaching the problem as follows:

1) Complete Stage I.
2) Move job A the minimum amount required to satisfy the constraint.
3) Evaluate net effects to determine if changes are required, but do not allow job A to become late.
4) Complete Stages II and III but do not consider moves which would make job A late.

None of the subjects achieved the cost which resulted from this approach. The strategy-trained subjects did generally begin with an approximation to Stage I, but their subsequent actions varied. One high analytic, strategy-trained subject (Transcript 3) described a decision process fairly close to the modified strategy.

There was no significant performance difference on the transfer task related to training treatment, although strategy-trained subjects performed 13% better on the average. Strategy-trained subjects did make more frequent use of both the net effect operator (p < .09 for a one-tailed Wilcoxon rank sum test) and the best net effect operator (p < .02), just as during the first decision.

There was also no significant performance difference related to CRA, although average performance
was approximately 20% higher for high analytic subjects. High analytic subjects requested net effect information significantly more frequently (p < .04 for a one-tailed Wilcoxon rank sum test).

While the strategy-trained group did not exhibit any significant performance advantage for the transfer task, the strategy training treatment was not designed to maximize the likelihood of transfer. Basic information concerning why and how the strategy worked was conveyed during training, but it was not stressed. Subjects were not told that they might be required to adapt the strategy to new situations, nor was the process of adapting the strategy ever modelled for them. Had these features been included in the training design (and more time available for training), the performance advantage might have transferred to the new situation. However, strategy training did meet a minimal criterion for generalization: subjects attempted to apply the strategy to the new task (Brown, 1978, p. 138).
6.9 SUMMARY OF THE FINDINGS

The most robust finding of this study was that strategy training led to higher performance than operations training. Substantial evidence suggests that strategy-trained subjects performed better because they behaved in a recognizably different manner from operations-trained subjects. The verbal protocols of strategy-trained subjects could be identified, and strategy-trained subjects devoted relatively fewer of their verbalizations to strategy formulation activities. The patterns of information use also differed by treatment and could be predicted with a fair degree of accuracy. Differences in information use and strategies were also apparent in self-reported information use and the user logs. The performance differences appeared to be related to strategy formulation rather than strategy execution. Far from having negative affective impacts, strategy-training appeared to lead to higher confidence and satisfaction with the decision strategy.

The performance advantage of strategy-trained subjects was sustained over time. While the delay was not long in terms of practical applications, the training and practice times for this experiment were
extremely brief in comparison to what might be expected in a practical application. The significant advantage of strategy-trained subjects did not generalize to a modified version of the task, but no specific actions were taken as part of the training design to enhance generalization of training.

There were some main effects of cognitive restructuring ability. For example, the statement-by-statement protocol analysis provided evidence that high analytics devoted more verbalizations to strategy formulation activities, especially method statements and observations. The most important finding regarding CRA, however, was that the relationship between CRA and performance depended on the training treatment.

Within the operations training treatment, high analytics performed better than low analytics. The higher performance of high analytics was traced back to differences in strategies. Global protocol analysis was used to establish that high analytics exhibited a more "analytic" decision approach. Retrospective, self-reported strategies, user log data and detailed, concurrent descriptions showed that operations-trained high and low analytics placed different importance on net effect information and developed decidedly
different methods for incorporating it into their decisions.

Within the strategy training treatment, low analytics actually outperformed high analytics slightly (although not significantly). While minor differences in perceptions and information use were detected, it is fair to say that the differences between ability groups were reduced along a number of dimensions through the provision of strategy training.

System designs need not be physically altered in order to respond to individual differences. As in this study, the users' state of knowledge concerning the system design can be altered instead.
### TABLE 11

**Descriptive Statistics for the Sample**

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Strategy</th>
<th>Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGE:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>23-44</td>
<td>27-44</td>
<td>23-44</td>
</tr>
<tr>
<td><strong>M</strong></td>
<td>31.5</td>
<td>32.6</td>
<td>30.3</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>5.2</td>
<td>4.7</td>
<td>5.5</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>39a</td>
<td>21</td>
<td>18a</td>
</tr>
<tr>
<td><strong>SEX:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td>27 (64%)</td>
<td>14 (67%)</td>
<td>13 (62%)</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td>15 (36%)</td>
<td>7 (33%)</td>
<td>8 (38%)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>42</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td><strong>GEFT:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>1-18</td>
<td>1-18</td>
<td>2-18</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td><strong>M</strong></td>
<td>11.8</td>
<td>11.9</td>
<td>11.7</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>5.3</td>
<td>5.3</td>
<td>5.5</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>42</td>
<td>21</td>
<td>21</td>
</tr>
</tbody>
</table>

*a Three subjects did not provide age information.

*b The theoretical range of scores on the GEFT is 0-18.
Figure 17. Distribution of GEFT scores across treatments.
**TABLE 12**

**Descriptive Statistics for the Protocol Group**

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Strategy</th>
<th>Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGE:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>26-44</td>
<td>29-44</td>
<td>26-44</td>
</tr>
<tr>
<td>M</td>
<td>36.2</td>
<td>37.2</td>
<td>34.8</td>
</tr>
<tr>
<td>SD</td>
<td>6.5</td>
<td>6.2</td>
<td>7.6</td>
</tr>
<tr>
<td>N</td>
<td>10(^a)</td>
<td>6</td>
<td>4(^a)</td>
</tr>
<tr>
<td><strong>SEX:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>4 (33%)</td>
<td>1 (17%)</td>
<td>3 (50%)</td>
</tr>
<tr>
<td>Female</td>
<td>8 (67%)</td>
<td>5 (83%)</td>
<td>3 (50%)</td>
</tr>
<tr>
<td>N</td>
<td>12</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td><strong>GEFT:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range(^b)</td>
<td>1-18</td>
<td>1-15</td>
<td>2-18</td>
</tr>
<tr>
<td>Median</td>
<td>10</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>M</td>
<td>9.5</td>
<td>9.3</td>
<td>9.7</td>
</tr>
<tr>
<td>SD</td>
<td>5.8</td>
<td>5.2</td>
<td>6.7</td>
</tr>
<tr>
<td>N</td>
<td>12</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

\(^a\)Two subjects did not provide age information.

\(^b\)The theoretical range of scores on the GEFT is 0-18.
**TABLE 13**

**GEFT Scores by Sex**

<table>
<thead>
<tr>
<th>GEFT</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>1-18</td>
<td>4-18</td>
</tr>
<tr>
<td>Median</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>(M)</td>
<td>8.9</td>
<td>13.4</td>
</tr>
<tr>
<td>(SD)</td>
<td>5.8</td>
<td>4.4</td>
</tr>
<tr>
<td>(N)</td>
<td>15</td>
<td>27</td>
</tr>
</tbody>
</table>
### TABLE 14

**Indices of the Reliability of Coding**

<table>
<thead>
<tr>
<th>Coding Category</th>
<th>Occur</th>
<th>Agree</th>
<th>Cohen’s Kappa</th>
<th>Z (Kappa)</th>
<th>Perreault &amp; Leigh’s I*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>50%</td>
<td>0.46</td>
<td>17.1</td>
<td>0.70</td>
</tr>
<tr>
<td>2</td>
<td>80</td>
<td>70%</td>
<td>0.68</td>
<td>22.2</td>
<td>0.83</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>60%</td>
<td>0.60</td>
<td>22.2</td>
<td>0.77</td>
</tr>
<tr>
<td>4</td>
<td>52</td>
<td>50%</td>
<td>0.48</td>
<td>18.1</td>
<td>0.70</td>
</tr>
<tr>
<td>5</td>
<td>219</td>
<td>70%</td>
<td>0.63</td>
<td>19.2</td>
<td>0.83</td>
</tr>
<tr>
<td>6</td>
<td>35</td>
<td>66%</td>
<td>0.65</td>
<td>22.3</td>
<td>0.80</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>36%</td>
<td>0.35</td>
<td>14.8</td>
<td>0.58</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>0%</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>0%</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>45%</td>
<td>0.44</td>
<td>17.3</td>
<td>0.66</td>
</tr>
<tr>
<td>11</td>
<td>151</td>
<td>75%</td>
<td>0.71</td>
<td>21.6</td>
<td>0.86</td>
</tr>
<tr>
<td>12</td>
<td>19</td>
<td>63%</td>
<td>0.63</td>
<td>22.3</td>
<td>0.79</td>
</tr>
<tr>
<td>13</td>
<td>78</td>
<td>58%</td>
<td>0.55</td>
<td>19.4</td>
<td>0.75</td>
</tr>
<tr>
<td>14</td>
<td>59</td>
<td>46%</td>
<td>0.44</td>
<td>17.0</td>
<td>0.66</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0%</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>4</td>
<td>50%</td>
<td>0.50</td>
<td>20.3</td>
<td>0.70</td>
</tr>
<tr>
<td>18</td>
<td>222</td>
<td>21%</td>
<td>0.13</td>
<td>5.8</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Overall 824 69% 0.65 48.1 0.82

95% Confidence
Lower Limit of I: 0.80

---

*Perreault & Leigh’s (1989) index of reliability was developed as an index of overall reliability. The index reported here for individual coding categories is a modified form of the original formula. The modified formula assumes that the ratio of unreliable codings which result in agreement is in proportion to the number of ways agreement could be achieved in comparison to the number of ways disagreement would result. For the case here of 17 categories, guesses were expected to result in agreement in one out of 33 cases (there are 33 possible outcomes involving a particular category, only one of which is an agreement).
Figure 18. Group cost performance means across job sets.
### TABLE 15

**Average Cost Performance for Job Set 1**

<table>
<thead>
<tr>
<th>Training Treatment</th>
<th>CRA</th>
<th>Row Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Strategy:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>93.5%</td>
<td>100%</td>
</tr>
<tr>
<td>$SD$</td>
<td>11.8%</td>
<td>0</td>
</tr>
<tr>
<td>$M_{GEFT}$</td>
<td>15.7</td>
<td>6.8</td>
</tr>
<tr>
<td>$N$</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Operations:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>87.0%</td>
<td>83.4%</td>
</tr>
<tr>
<td>$SD$</td>
<td>15.1%</td>
<td>18.9%</td>
</tr>
<tr>
<td>$M_{GEFT}$</td>
<td>16.2</td>
<td>6.8</td>
</tr>
<tr>
<td>$N$</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Column Statistics:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>90.4%</td>
<td>91.3%</td>
</tr>
<tr>
<td>$SD$</td>
<td>13.6%</td>
<td>15.9%</td>
</tr>
<tr>
<td>$M_{GEFT}$</td>
<td>15.9</td>
<td>6.8</td>
</tr>
<tr>
<td>$N$</td>
<td>23</td>
<td>19</td>
</tr>
</tbody>
</table>
### TABLE 16

**Average Cost Performance for Job Set 2**

<table>
<thead>
<tr>
<th>Training Treatment</th>
<th>High</th>
<th>Low</th>
<th>Row Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CRA</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategy:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>90.0%</td>
<td>93.7%</td>
<td>91.5%</td>
</tr>
<tr>
<td>$SD$</td>
<td>12.5%</td>
<td>13.0%</td>
<td>12.4%</td>
</tr>
<tr>
<td>$M_{GEFT}$</td>
<td>15.9</td>
<td>5.8</td>
<td>11.9</td>
</tr>
<tr>
<td>$N$</td>
<td>9</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Operations:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>78.2%</td>
<td>66.7%</td>
<td>72.7%</td>
</tr>
<tr>
<td>$SD$</td>
<td>14.0%</td>
<td>22.8%</td>
<td>19.1%</td>
</tr>
<tr>
<td>$M_{GEFT}$</td>
<td>16.2</td>
<td>6.8</td>
<td>11.7</td>
</tr>
<tr>
<td>$N$</td>
<td>11</td>
<td>10</td>
<td>21</td>
</tr>
<tr>
<td>Column Statistics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>83.5%</td>
<td>76.8%</td>
<td>80.5%</td>
</tr>
<tr>
<td>$SD$</td>
<td>14.3%</td>
<td>23.0%</td>
<td>19.0%</td>
</tr>
<tr>
<td>$M_{GEFT}$</td>
<td>16.1</td>
<td>6.4</td>
<td>11.8</td>
</tr>
<tr>
<td>$N$</td>
<td>20</td>
<td>16</td>
<td>36</td>
</tr>
</tbody>
</table>
TABLE 17

Average Cost Performance for Job Set 3

<table>
<thead>
<tr>
<th>Training Treatment</th>
<th>CRA</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Strategy:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>87.5%</td>
<td>98.1%</td>
<td>89.8%</td>
</tr>
<tr>
<td>SD</td>
<td>19.0%</td>
<td>0</td>
<td>17.1%</td>
</tr>
<tr>
<td>M_{GEFT}</td>
<td>15.4</td>
<td>2.5</td>
<td>12.6</td>
</tr>
<tr>
<td>N</td>
<td>7</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Operations:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>69.2%</td>
<td>50.3%</td>
<td>59.7%</td>
</tr>
<tr>
<td>SD</td>
<td>18.0%</td>
<td>21.4%</td>
<td>21.6%</td>
</tr>
<tr>
<td>M_{GEFT}</td>
<td>16.2</td>
<td>6.8</td>
<td>11.5</td>
</tr>
<tr>
<td>N</td>
<td>10</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Column Statistics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>76.7%</td>
<td>58.2%</td>
<td>69.1%</td>
</tr>
<tr>
<td>SD</td>
<td>20.1%</td>
<td>26.9%</td>
<td>24.5%</td>
</tr>
<tr>
<td>M_{GEFT}</td>
<td>15.9</td>
<td>6.1</td>
<td>11.8</td>
</tr>
<tr>
<td>N</td>
<td>17</td>
<td>12</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>JOB SET 1</td>
<td>JOB SET 2</td>
<td>JOB SET 3</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>R-square</td>
<td>.256</td>
<td>.322</td>
<td>.429</td>
</tr>
<tr>
<td>Pr &gt; F</td>
<td>p &lt; .01</td>
<td>p &lt; .006</td>
<td>p &lt; .003</td>
</tr>
<tr>
<td><strong>Main Effect:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>p &lt; .004</td>
<td>p &lt; .009</td>
<td>p &lt; .007</td>
</tr>
<tr>
<td><strong>Main Effect:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEFT</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Interaction Effect:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment*GEFT</td>
<td>p &lt; .04</td>
<td>p &lt; .14</td>
<td>p &lt; .135</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>42</td>
<td>36</td>
<td>29</td>
</tr>
</tbody>
</table>
Figure 19. Graphs of cost performance regression lines.
### TABLE 19

**Analysis of Contrasts within the Cost Performance Data**

<table>
<thead>
<tr>
<th>Prediction</th>
<th>JOB SET 1</th>
<th>JOB SET 2</th>
<th>JOB SET 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HS = LS:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction</td>
<td>LS &gt; HS</td>
<td>LS &gt; HS</td>
<td>LS &gt; HS</td>
</tr>
<tr>
<td>Significance Level$^a$</td>
<td>p &lt; .12</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>HO &gt; LO:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction</td>
<td>HO &gt; LO</td>
<td>HO &gt; LO</td>
<td>HO &gt; LO</td>
</tr>
<tr>
<td>Significance Level$^b$</td>
<td>n.s.</td>
<td>p &lt; .09</td>
<td>p &lt; .03</td>
</tr>
<tr>
<td><strong>HS &gt; HO:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction</td>
<td>HS &gt; HO</td>
<td>HS &gt; HO</td>
<td>HS &gt; HO</td>
</tr>
<tr>
<td>Significance Level$^b$</td>
<td>n.s.</td>
<td>p &lt; .04</td>
<td>p &lt; .04</td>
</tr>
</tbody>
</table>

$^a$The reported significance levels are for two-tailed t-tests on the subgroup means.

$^b$The reported significance levels are for one-tailed t-tests on the subgroup means.
Figure 20. Group time performance means across job sets.
### TABLE 20

**Average Time Performance (in Seconds) for Job Set 1**

<table>
<thead>
<tr>
<th>Training Treatment</th>
<th>CRA</th>
<th></th>
<th></th>
<th>Row Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>High</strong></td>
<td><strong>Low</strong></td>
<td><strong>Mean</strong></td>
<td><strong>Statistics</strong></td>
</tr>
<tr>
<td>Strategy:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>1410</td>
<td>2179</td>
<td>1740</td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td>933</td>
<td>1061</td>
<td>1040</td>
<td></td>
</tr>
<tr>
<td>$M_{GEFT}$</td>
<td>15.7</td>
<td>6.8</td>
<td>11.9</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>12</td>
<td>9</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Operations:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>1049</td>
<td>1298</td>
<td>1161</td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td>640</td>
<td>605</td>
<td>621</td>
<td></td>
</tr>
<tr>
<td>$M_{GEFT}$</td>
<td>16.2</td>
<td>6.9</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>11</td>
<td>9</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Column Statistics:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>1237</td>
<td>1739</td>
<td>1458</td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td>810</td>
<td>952</td>
<td>900</td>
<td></td>
</tr>
<tr>
<td>$M_{GEFT}$</td>
<td>15.9</td>
<td>6.8</td>
<td>11.9</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>23</td>
<td>18</td>
<td>41</td>
<td></td>
</tr>
</tbody>
</table>
**TABLE 21**

**ANOVA Results for Time Performance Data**

<table>
<thead>
<tr>
<th></th>
<th>JOB SET 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R)-square</td>
<td>( .209 )</td>
</tr>
<tr>
<td>( Pr &gt; F )</td>
<td>( p &lt; .04 )</td>
</tr>
<tr>
<td><strong>Main Effect:</strong></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>( p &lt; .03 )</td>
</tr>
<tr>
<td>CRA</td>
<td>( p &lt; .06 )</td>
</tr>
<tr>
<td><strong>Interaction</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Effect:</strong></td>
<td></td>
</tr>
<tr>
<td>Treatment*CRA</td>
<td>( \text{n.s.} )</td>
</tr>
</tbody>
</table>

\( N \)             \( 41 \)
**TABLE 22** (continued on next page)

**Associations between Information Item Use and Training Treatment**

<table>
<thead>
<tr>
<th>Information Item</th>
<th>Predicted</th>
<th>Strength of Association</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Schedule Comparison Screen:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Sequences:</td>
<td>O &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td>Subtotals by Cost Category:</td>
<td>O &gt; S</td>
<td>p &lt; .02</td>
</tr>
<tr>
<td>Totals by Sequencing Rule:</td>
<td>O &gt; S</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Modification Screen:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changeover Charges Column:</td>
<td>S &gt; O</td>
<td>n.s.</td>
</tr>
<tr>
<td>Tardiness Charges Column:</td>
<td>O &gt; S</td>
<td>p &lt; .0005</td>
</tr>
<tr>
<td>Inventory Charges Column:</td>
<td>O &gt; S</td>
<td>p &lt; .0005</td>
</tr>
<tr>
<td>Period Completed Column:</td>
<td>O &gt; S</td>
<td>p &lt; .003</td>
</tr>
<tr>
<td>Subtotals by Cost Category:</td>
<td>O &gt; S</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

**Note.** The strength of the association was tested by comparing the proportion of subjects who reported using each item. The significance level is for a chi square test of association. All of the associations were in the expected direction. Significance levels reported for tests on the number of items used are for one-tailed Wilcoxon rank sum tests. S = Strategy-trained, O = Operations-trained.
TABLE 22 (continued from previous page)

**Associations between Information Item Use and Training Treatment**

<table>
<thead>
<tr>
<th>Information Item</th>
<th>Predicted</th>
<th>Strength of Association</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analysis Screen:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changeover Charges Column:</td>
<td>0 &gt; S</td>
<td>p &lt; .04</td>
</tr>
<tr>
<td>Inventory Costs per Period:</td>
<td>0 &gt; S</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td>Periods Processing Required:</td>
<td>0 &gt; S</td>
<td>p &lt; .0005</td>
</tr>
<tr>
<td>Tardiness Costs per Period:</td>
<td>0 &gt; S</td>
<td>p &lt; .0005</td>
</tr>
<tr>
<td>Periods Late:</td>
<td>0 &gt; S</td>
<td>p &lt; .0005</td>
</tr>
<tr>
<td>Net Effect:</td>
<td>S &gt; O</td>
<td>n.s.</td>
</tr>
<tr>
<td>Best Net Effect:</td>
<td>S &gt; O</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

**Note.** The strength of the association was tested by comparing the proportion of subjects who reported using each item. The significance level is for a chi square test of association. All of the associations were in the expected direction. Significance levels reported for tests on the number of items used are for one-tailed Wilcoxon rank sum tests. S = Strategy-trained, O = Operations-trained.
TABLE 22 (continued from previous page)

**Associations between Information Item Use and Training Treatment**

<table>
<thead>
<tr>
<th>Information Item</th>
<th>Predicted</th>
<th>Strength of Association</th>
</tr>
</thead>
</table>

**High/Low Values Screen:**

- Number of Items Used: **0 > S** \( p < .0003 \)

**Changeover Costs Screen:**

- **0 > S** \( \text{n.s.} \)

**Total Number of Information Items Used:**

- **0 > S** \( p < .0001 \)

**Note.** The strength of the association was tested by comparing the proportion of subjects who reported using each item. The significance level is for a chi square test of association. All of the associations were in the expected direction. Significance levels reported for tests on the number of items used are for one-tailed Wilcoxon rank sum tests. **S** = Strategy-trained, **O** = Operations-trained.


**TABLE 23**  

**Associations Between Information Item Use and CRA within Training Treatments**

<table>
<thead>
<tr>
<th>Information Item</th>
<th>Predicted</th>
<th>Strength of Association</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strategy Training Treatment:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Modification Screen:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotals by Cost Category:</td>
<td>( H &gt; L )</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Analysis Screen:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Effect:</td>
<td>( H &gt; L )</td>
<td>( p &lt; .09 )</td>
</tr>
<tr>
<td>Best Net Effect:</td>
<td>( H &gt; L )</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Changeover Costs Screen:</strong></td>
<td>( H &gt; L )</td>
<td>( p &lt; .07 )</td>
</tr>
<tr>
<td><strong>Total Number of Information Items Used:</strong></td>
<td>( L &gt; H )</td>
<td>n.s.</td>
</tr>
<tr>
<td>(wrong direction)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Operations Training Treatment:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Schedule Comparison Screen:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Sequences:</td>
<td>( H = L )</td>
<td>( p &lt; .10 ), ( L &gt; H )</td>
</tr>
<tr>
<td>Totals by Sequencing Rule:</td>
<td>( H = L )</td>
<td>( p &lt; .07 ), ( L &gt; H )</td>
</tr>
</tbody>
</table>

**Note.** The strength of the association was tested by comparing the proportion of subjects who reported using each item. Because at least one of the expected cell frequencies was less than five in each case, a Fisher’s exact test was used to determine the significance of the association. \( H \) = High CRA. \( L \) = Low CRA.
TABLE 24

Results of Aggregate Analysis of Protocols for Treatment Effect

<table>
<thead>
<tr>
<th>Transcript Number</th>
<th>Actual</th>
<th>GEFT Score</th>
<th>Assigned Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rater 1</td>
</tr>
<tr>
<td>1</td>
<td>S</td>
<td>11</td>
<td>S</td>
</tr>
<tr>
<td>2</td>
<td>O</td>
<td>17</td>
<td>O</td>
</tr>
<tr>
<td>3</td>
<td>S</td>
<td>14</td>
<td>S</td>
</tr>
<tr>
<td>4</td>
<td>S</td>
<td>6</td>
<td>S</td>
</tr>
<tr>
<td>5</td>
<td>O</td>
<td>18</td>
<td>O</td>
</tr>
<tr>
<td>6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>S</td>
<td>9</td>
<td>S</td>
</tr>
<tr>
<td>7</td>
<td>S</td>
<td>15</td>
<td>S</td>
</tr>
<tr>
<td>8</td>
<td>O</td>
<td>5</td>
<td>O</td>
</tr>
<tr>
<td>9</td>
<td>O</td>
<td>5</td>
<td>O</td>
</tr>
<tr>
<td>10</td>
<td>O</td>
<td>2</td>
<td>O</td>
</tr>
<tr>
<td>11</td>
<td>S</td>
<td>1</td>
<td>S</td>
</tr>
<tr>
<td>12</td>
<td>O</td>
<td>11</td>
<td>O</td>
</tr>
</tbody>
</table>

Misclassifications: None 3


<sup>a</sup>This subject was absent the day of the lecture on job scheduling.
### TABLE 25

**Results of the Statement-by-Statement Protocol Analysis**

<table>
<thead>
<tr>
<th>Coding Category</th>
<th>High</th>
<th>Low</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Method Statement</td>
<td>16%</td>
<td>10%</td>
<td>p = .10</td>
</tr>
<tr>
<td>2) Specific Move Description</td>
<td>6%</td>
<td>9%</td>
<td>n.s.</td>
</tr>
<tr>
<td>4) Observation</td>
<td>9%</td>
<td>5%</td>
<td>p = .05</td>
</tr>
<tr>
<td>5) Information Selection</td>
<td>15%</td>
<td>13%</td>
<td>n.s.</td>
</tr>
<tr>
<td>6) Prediction/Knowledge of System Behavior</td>
<td>5%</td>
<td>1%</td>
<td>n.s.</td>
</tr>
<tr>
<td>1-9) Sum of Analytic Coding Incidents</td>
<td>54%</td>
<td>38%</td>
<td>p = .10</td>
</tr>
</tbody>
</table>

*Note.* Significance levels are reported for one-tailed Wilcoxon rank sum tests using 12 subjects (4 high analytics, 8 low analytics).
### TABLE 26 (continued on next page)

**Summary of Protocol Subjects' Methods for Incorporating Net Effect Information into the Decision**

<table>
<thead>
<tr>
<th>Transcript</th>
<th>GEFT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>17</td>
</tr>
</tbody>
</table>

**Method:** "This was most useful in finding the best cost...Made changes [swaps] based on - net effects or small + net effects" (questionnaire response and observer notes). "The net effect was the inventory holding cost and the tardiness charges if you were to change it. It’s the difference between the 2 being changed. Basically, there are four numbers involved in the calculation...if you switched it, how that affected between those 4" (from follow-up questioning).

| 5          | 18   |

**Method:** "I just keep going until I don’t find any negatives and do [swap the pairs associated with] the net effects" (transcript line 1 and observer notes).

| 8          | 5    |

**Method:** Did not use: "I never grasped what the net effects information was used for...or their significance" (questionnaire response).
<table>
<thead>
<tr>
<th>Transcript</th>
<th>GEFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td><strong>Method:</strong> &quot;Only in an attempt to correlate if a change occur&quot; (questionnaire response). &quot;It's more quantitative to do it this way but I'm not really sure what I'm looking at here&quot; (transcript line 28).</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td><strong>Method:</strong> &quot;I'm trying to put all of the negatives negative net effects together (transcript line 197). &quot;A and C...negative--best computed net effect I&amp;T net effect. Hell, I don't even know what that means&quot; (transcript lines 156-157).</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td><strong>Method:</strong> &quot;I'm not even sure I can tell you exactly what the theory is other than I know that those are the net effect numbers that have the highest impact&quot; (quote from follow-up questioning). &quot;I move one up or down [by no fixed amount] depending on which way I move them&quot; (quote from follow-up questioning and observer notes).</td>
<td></td>
</tr>
<tr>
<td>Regression Analysis Results for Net Effect Operator Use</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>JOB SET 1</strong></td>
<td><strong>JOB SET 2</strong></td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>R-square</strong></td>
<td>.216</td>
</tr>
<tr>
<td><strong>Pr &gt; F</strong></td>
<td>p &lt; .04</td>
</tr>
<tr>
<td><strong>Main Effect:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td>p &lt; .08</td>
</tr>
<tr>
<td><strong>GEFT</strong></td>
<td>p &lt; .14</td>
</tr>
<tr>
<td><strong>Interaction Effect:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Treatment*GEFT</strong></td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>38</td>
</tr>
</tbody>
</table>
### TABLE 28

**Analysis of Contrasts within the Net Effect Operator Use Data**

<table>
<thead>
<tr>
<th>Prediction</th>
<th>JOB SET 1</th>
<th>JOB SET 2</th>
<th>JOB SET 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HS = LS:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction</td>
<td>HS &gt; LS</td>
<td>HS &gt; LS</td>
<td>HS &gt; LS</td>
</tr>
<tr>
<td>Significance Level(^a)</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>HO &gt; LO:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direction</td>
<td>HO &gt; LO</td>
<td>HO &gt; LO</td>
<td>HO &gt; LO</td>
</tr>
<tr>
<td>Significance Level(^b)</td>
<td>p &lt; .04</td>
<td>p &lt; .004</td>
<td>p &lt; .003</td>
</tr>
</tbody>
</table>

\(^a\)The reported significance levels are for two-tailed Wilcoxon rank sum tests on the subgroup means.

\(^b\)The reported significance levels are for one-tailed Wilcoxon rank sum tests on the subgroup means.
### TABLE 29

**Analysis of Contrasts within the Best Net Effect**

**Operator Use Data**

<table>
<thead>
<tr>
<th>Prediction</th>
<th>JOB SET 1</th>
<th>JOB SET 2</th>
<th>JOB SET 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S &gt; 0:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Direction</strong></td>
<td>S &gt; 0</td>
<td>S &gt; 0</td>
<td>S &gt; 0</td>
</tr>
<tr>
<td><strong>Significance Level</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>p &lt; .0003</td>
<td>p &lt; .0001</td>
<td>p &lt; .0009</td>
</tr>
<tr>
<td><strong>H &gt; L:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Direction</strong></td>
<td>H &gt; L</td>
<td>H &gt; L</td>
<td>H &gt; L</td>
</tr>
<tr>
<td><strong>Significance Level</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>n.s.</td>
<td>p &lt; .06</td>
<td>p &lt; .003</td>
</tr>
<tr>
<td><strong>HS = LS:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Direction</strong></td>
<td>LS &gt; HS</td>
<td>LS &gt; HS</td>
<td>HS &gt; LS</td>
</tr>
<tr>
<td><strong>Significance Level</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>HO &gt; LO:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Direction</strong></td>
<td>HO &gt; LO</td>
<td>HO &gt; LO</td>
<td>HO &gt; LO</td>
</tr>
<tr>
<td><strong>Significance Level</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>n.s.</td>
<td>p &lt; .005</td>
<td>p &lt; .003</td>
</tr>
</tbody>
</table>

<sup>a</sup>The reported significance levels are for two-tailed Wilcoxon rank sum tests on the subgroup means.

<sup>b</sup>The reported significance levels are for one-tailed Wilcoxon rank sum tests on the subgroup means.
### TABLE 30

Regression Analysis Results for Calculations Performed

<table>
<thead>
<tr>
<th></th>
<th>JOB SET 1</th>
<th>JOB SET 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R-square</strong></td>
<td>.214</td>
<td>.392</td>
</tr>
<tr>
<td><strong>Pr &gt; F</strong></td>
<td>p &lt; .04</td>
<td>p &lt; .002</td>
</tr>
<tr>
<td><strong>Main Effect:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>n.s.</td>
<td>p &lt; .06</td>
</tr>
<tr>
<td><strong>Main Effect:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEFT</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Interaction Effect:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment*GEFT</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>38</td>
<td>36</td>
</tr>
</tbody>
</table>

\(^a\)The treatment effect was significant when the interaction term was removed from the model (p < .009).
TABLE 31

Average Overall Satisfaction with Strategy

<table>
<thead>
<tr>
<th>Training Treatment</th>
<th>CRA</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategy:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>5.4</td>
<td>4.9</td>
<td>5.2</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.2</td>
<td>1.9</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>9</td>
<td>9</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Operations:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>4.8</td>
<td>2.4</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.5</td>
<td>0.9</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>11</td>
<td>9</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Column Statistics:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>5.1</td>
<td>3.7</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.4</td>
<td>1.9</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>18</td>
<td>38</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 7. CONCLUSION

This chapter presents the contributions and limitations of the research. The final section discusses the practical applicability of strategy training.

7.1 CONTRIBUTIONS AND LIMITATIONS

The most important contribution of this work is to illustrate that system designs do not have to be altered in order to respond to individual differences. Undoubtedly, individual differences will manifest themselves as strategy differences when users are forced to adapt to a system with no assistance or direction. However, the strategies which individuals spontaneously devise do not indicate that they are incapable of adopting a different approach or unwilling to do so.

Although CRA was selected for this study, there is no reason to believe that the results are peculiar to this individual difference. A vast array of individual differences combine to determine the nature of the strategy a user creates. The strategy determines the user's resulting performance and the user's knowledge and support requirements. As long as users are willing
and able to understand and execute a suggested strategy, strategy training leads to more uniform strategies and more uniform support requirements. The result is the same regardless of the origin of differences in spontaneously devised strategies or the nature of those differences. This is the advantage of addressing individual differences through training rather than design: I do not need to predict or differences in behavior or preferences because I have the power to influence behavior and preferences if I choose to exercise it.

Future research needs to explore the limits to which users are "willing and able to understand and execute a suggested strategy." The limits of a training approach might be related to individual differences. Within this study, there was some evidence that CRA was associated with an individual's willingness to adopt the suggested strategy. On the exit questionnaire, subjects indicated why they had not completely followed the suggested strategy. For low analytics, deviations were either unintentional or took the form of additional error checking. On the contrary, three high analytics intentionally made substantive changes to the suggested strategy (and contributed to the lower performance of HS subjects).
This result is consistent with the finding that field dependents are more receptive to an "authoritative" view (Witkin et al, 1971, p. 9; Witkin et al, 1977, pp. 10-11).

The extent to which users are willing and able to execute suggested strategies will also certainly depend on the nature of the suggested strategy. One of the limitations of this study is that the suggested strategy is programmable. Given the limited amount of time available for training, the nature of the suggested strategy was inadequate in order to experimentally establish that the training approach was a possible means of responding to individual differences.

As the domain becomes less structured, I would expect that the amount of time required for training would increase and the amount of detail provided during training would decrease. However, I would also expect that some guidance is preferable to no guidance. Even the guidance provided by an internalized checklist of activities might improve performance (Brown, 1978, p. 157). Strategy training in less structured situations might focus on appropriate methods without any specification of goals, or vice versa. Cats-Baril & Huber (1987) found performance improvements associated
with provision of a heuristic to solve an ill-structured problem.

Another limitation of this study was that the suggested strategy did not have to compete with an existing decision process, a situation which would often occur in practice. Where prior procedures already exist, the success of adoption would be far more dependent on the credibility of the advocates of the suggested strategy, the strength of the arguments in its favor, and the evidence that it leads to superior performance. Existing procedures may not prove to be much of an obstacle to strategy training if, as in the current study, the users are not very satisfied with their existing methods.

Another contribution of this research is to clarify that the GEFT is an ability test. While CRA is related to the cognitive style dimension of field independence/dependence, it is not equivalent to it. A proper conceptualization of CRA is important in identifying relevant literature and in understanding how CRA is related to behavior and performance.

A third contribution of this study is to establish that training can substantially enhance performance. The superiority of strategy training in this study parallels the result found by Hellman (1989, p. 121):
"the use situation of an information system can be improved by improving the users' qualifications alone, without modifying the technical system or the user interface." As discussed in the literature review, past training studies have provided remarkably little tangible evidence that resources devoted to training are well spent. This research suggests that strategy training provides substantial performance gains over operations training.

Finally, it is my hope that future MIS/DSS research concerning training will at least meet the standard of explanation provided in this work. When alternative training treatments are selected for study, there should be a clear explication as to why and how the learning outcomes are expected to differ. The method of testing outcomes should reflect the nature of expected differences and explicitly consider the extent to which generalization is being required of the subjects. The GOMS framework is a useful means of summarizing alternative forms of training and testing.

7.2 DISCUSSION

This research clearly established that a strategy training approach was effective in the laboratory. Further research is required in order to determine
under what circumstances strategy training can and should be implemented in practice. Even without additional research, however, a number of points regarding motives and opportunity can be made.

First, without claiming that strategy training is always appropriate, there are practical opportunities to apply the method. The existence of a suggested strategy or normative approach need not be "antithetical to a DSS approach" (Silver, 1987, p. 13). For example, Alter's (1977) findings suggest that strategy training is particularly important for data analysis systems:

...a recurrent problem was that system implementers and proponents incorrectly assumed that potential users would figure out how to apply the systems; in the more successful cases, either users were trained to use the system in a relatively repetitive manner or the implementers themselves were the users.

The Portfolio Management System, as described in Keen & Scott Morton (1978, pp. 99-126), provides a specific illustration of both the need for strategy training and potential problems of user acceptance of training. More currently, a systems analyst who took part in this study indicated that a lack of understanding of system
outputs was a very real problem in her organization which might be resolved through strategy training.

Strategy training is a natural extension of a decision-oriented approach to design, as described by Stabell (1983, p. 258). The motivations for adopting a decision-oriented approach are necessarily linked to larger issues of change agency as discussed by Silver (1989). It is not unusual for the designers of a DSS to act as agents of directed change, "since a major motivation in [DSS] development is the desire by the client for an improvement in the user's planning process" (Moore & Chang, 1983, p. 189).

There is no question that a decision-oriented approach emphasizing strategy training entails a greater investment throughout the system life cycle. In situations in which the decision is not repeated or only a single, skilled user makes use of the system, the costs would probably not be offset by the benefits. However, the learning and performance benefits of employing strategy training for repetitive decisions are likely to exceed the additional costs. As in this study, the outcome of the design process might even be a decision which could be automated.

The main risk associated with a strategy training approach is that the suggested strategy might be less
effective than the majority of existing strategies. We can partially control that risk by selecting the individuals who devise the suggested strategy, presumably on the basis of superior observed skill and qualifications. We can give those individuals time in which to create the strategy, draw upon the knowledge and insights of others, and test alternative approaches.

In assessing the risk of a suggested strategy, we must also consider the comparative quality of strategies developed outside of the partially controlled conditions described above. The quality of those strategies will vary with the ability of the user, but there are reasons why even very talented individuals might formulate poor strategies for using a system. First, the situation in which an individual must adapt with minimal guidance is the most demanding of learning situations. Learning from experience alone may mean not learning at all (Brehmer, 1980). Second, if discovery learning is to be effective, there must be time in which to make discoveries. Unfortunately, uninterrupted time in which to reflect on a decision process is not characteristic of the nature of managerial work (Mintzberg, 1973). Third, a user with the ability to devise good strategies may not be
motivated to do so in a particular decision context. Finally, an otherwise brilliant decision maker may lack specific domain knowledge and experience (e.g., a new employee).

In designing a system, we must determine which risk we find more acceptable: the risk associated with a single, known decision process developed under partially controlled conditions or the risk associated with multiple, unknown decision processes developed under uncontrolled conditions. By calling attention to the range of skill in problem-solving even within a group of MBA students, this study illustrates that the risks associated with a nondirective approach are considerable. Stabell (1983, p. 231) goes so far as to suggest that "managers alone and unaided are not likely to be able to systematically improve their decision making."

Within a decision-oriented approach, decision processes can be influenced indirectly through the system design and/or more directly through strategy training. Given the potential for differences in interpretation of any system output, providing strategy training may be the safer route. There is no guarantee that the user, alone and unaided, will come to understand the system as the designer intended. For
example, some of the operations-trained subjects in this study interpreted the net effect information in a manner which I had certainly not intended.

Interestingly, arguments in favor of a nondirective, user-controlled system interaction have their parallel in the instructional literature in the form of arguments advocating learner-controlled instruction. However, instruction in accordance with learner preferences has not been found to enhance learning and may be detrimental (Cronbach & Snow, 1977, p. 170). Atkinson (1972, p. 930) was "alarmed" by the number of individuals advocating completely learner-controlled instruction:

My data, and the data of others, indicate that the learner is not a particularly effective decision maker. Arguments against learner-controlled programs are unpopular in the present climate of opinion, but they need to be made so that one will not be seduced by the easy answer that a theory of instruction is not required because "who can be a better judge of what is best for the student than the student himself."

Similarly, arguments for flexible and adaptive systems provide an easy answer to the problem of decision support (Stabell, 1983, p. 224).

I suggested in the introduction to this work that the boundaries between the concepts of adaptation,
learning and problem-solving are virtually nonexistent. If learner control does not generally lead to improved learning performance, why would decision maker control of MIS/DSS lead to improved decision-making performance? While I may be in the minority in resisting the allure of flexibility and user control in favor of strategy training, I am also in very illustrious company. I dare claim to move in the same circles as the folk who decided that the wheel was a good enough idea to be shared, even with the people who had not invented it. You can judge for yourself whether strategy training works outside of the laboratory.
Appendix 1.

Operating Help Screens
OPENING SCHEDULE COMPARISON SCREEN

AFTER VIEWING the comparison screen, PRESS "M" to MODIFY a schedule. You will then be able to make changes to the schedule with the lowest initial cost. The schedule comparison screen is displayed each time you begin a new job set. It displays the results (total cost and cost by category) of applying 5 sequencing rules to the new jobs. The 5 sequencing rules are as follows:

SMITH'S RULE: Jobs appear in increasing order of
(processing time) / (inventory holding cost per period)

SHORTEST PROCESSING TIME RULE: Jobs appear in increasing order of
processing time.

WEIGHTED SLACKNESS RULE: Jobs appear in increasing order of
(slack) / (tardiness cost per period)

LEAST CHANGEOVER HEURISTIC: Jobs appear in an order which minimizes the first changeover cost and all subsequent pairwise costs.

RANDOM ORDER: Jobs are arranged alphabetically from A to G.
PgUp — Previous FgDn — Next
Help Screen
            End — Leave Help
            Help Screen
ABOUT THE INFORMATION DISPLAYED ON THE SCREEN:

Along with the total cost information for the schedule currently displayed on the screen, you are also provided with total cost information for the previous schedule (this will only exist if you have made at least one change) and a number called "BEST ACHIEVED." This number represents the lowest total cost schedule you have achieved in the course of modifying the schedule displayed on your screen.

TO VIEW ADDITIONAL INFORMATION BEFORE MAKING A CHANGE:

Press "A" for ANALYZE. You will be provided with information concerning the duration of each job, changeover charges, inventory holding costs per period, tardiness costs per period, and job lateness. You will also have access to the Schedule Comparison Screen, a Changeover Costs Matrix, and a screen displaying the High and Low Values of various job characteristics.

OTHER: Press PgDn for the next help screen.
PgUp -- Previous   PgDn -- Next   End -- Leave Help
Help Screen        Help Screen
MODIFICATION SCREEN (continued)

TO MAKE A CHANGE:

1) Use the UP and DOWN arrow keys to highlight the job you wish moved.

* 2) Use the LEFT arrow key to mark the job to be moved. If you mark the wrong job, use the DEL key to remove the marker.

3) Use the up and down arrow keys to highlight the new position in which the marked job should be placed.

* 4) Use the RIGHT arrow key to mark the new location. At this point, if either of the locations marked is incorrect, use the DEL key to remove the marker. Press DEL once and the location marker is erased. Press DEL again and the move job marker is erased.

* 5) Once the job to be moved and its new location are marked, press the ENTER key to COMPUTE the costs of the revised schedule.

OTHER: Press PgDn for the next help screen.

PgUp -- Previous                  PgDn -- Next                  End -- Leave Help
       Help Screen                  Help Screen
MODIFICATION SCREEN (continued)

TO "UNDO" A CHANGE:

To undo the very last change that you made, press "P" to restore the immediately PREVIOUS version of the schedule.

To undo a series of changes that you have made since achieving the lowest total cost version of this schedule, press "B" to restore the BEST version which you have created so far for this schedule.

The DEL key removes a left/right arrow marking a job or location.

TO INDICATE THAT YOU HAVE FINISHED MAKING CHANGES:

If you are COMPLETELY finished making changes to the schedule on your screen, press "F" for "FINISHED Making Changes." You will not be able to return to the current schedule later, so make sure that you ARE finished.

PgUp -- Previous           PgDn -- Next            End -- Leave Help
Help Screen               Help Screen
TO MAKE A CHANGE or TO INDICATE THAT YOU HAVE FINISHED MAKING CHANGES:

Press "M" to Return to the MODIFICATION Screen. Once there, you can either make a change, modify another schedule or proceed to the next job set.

INVENTORY AND TARDINESS (I&T) NET EFFECT:

The system can automatically perform a calculation which considers simultaneously the inventory holding costs and tardiness costs for any 2 jobs highlighted on the analysis screen. The resulting number, called the I&T Net Effect, shows the overall effect on these costs if the positions of the 2 jobs were reversed. For example, a net effect of -100 means that the sum of inventory and tardiness charges would be reduced by 100 if the order of the 2 jobs were reversed. THIS NUMBER DOES NOT TAKE INTO ACCOUNT THE EFFECT ON CHANGEOVER COSTS.

FOR MORE INFORMATION CONCERNING NET EFFECT or FOR OTHER INFORMATION:
Press PgDn for next help screen.
PgUp -- Previous PgDn -- Next End -- Leave Help
Help Screen Help Screen
ANALYSIS (continued)

TO CALCULATE THE I&T NET EFFECT FOR A PAIR OF JOBS:

Use the up and down arrow keys to highlight the desired 2 jobs (only adjacent jobs can be highlighted). Then press "E" for "EVALUATE Net Effect." The result of the calculation is displayed in the upper right of the screen.

BEST NET EFFECT:

The system automatically monitors the lowest net effect which you have evaluated so far and displays that number on the right of the screen. At any time that you wish to highlight the 2 jobs which produced the lowest net effect, press "B" to highlight the BEST Net Effect.

TO VIEW ADDITIONAL INFORMATION CONCERNING THE JOB SET:

Press PgDn for the next help screen.

PgUp -- Previous Help Screen
PgDn -- Next Help Screen
End -- Leave Help
TO VIEW ADDITIONAL INFORMATION CONCERNING THE JOB SET:

Press "C" to view the Schedule COMPARISON Screen. For further information concerning the Schedule Comparison Screen, press PgDn.

Press "T" to display a Changeover Costs TABLE (Matrix).

Press "H" to display a HIGH and Low Values Screen which lists the highest/lowest values and corresponding jobs for inventory holding costs, tardiness costs, slack, duration, changeover costs, the ratio used in Smith's Rule and the ratio used in the Weighted Slackness rule.
SCHEDULE COMPARISON SCREEN

The schedule comparison screen displays 5 sequencing rules which have been applied to the jobs in the current job set. The 5 rules are as follows:

**SMITH'S RULE:** Jobs appear in increasing order of 
(processing time) / (inventory holding cost per period)

**SHORTEST PROCESSING TIME RULE:** Jobs appear in increasing order of 
processing time.

**WEIGHTED SLACKNESS RULE:** Jobs appear in increasing order of 
(slack) / (tardiness cost per period)

**LEAST CHANGEOVER HEURISTIC:** Jobs appear in an order which minimizes the first changeover cost and all subsequent pairwise costs.

**RANDOM ORDER:** Jobs are arranged alphabetically from A to G.

PgUp -- Previous 
PgDn -- Next 
Help Screen 
End -- Leave Help 
Help Screen
Appendix 2.

Strategy Help Screens
The system automatically selects the schedule with the initial LOWEST TOTAL COST as the starting point for modifications. Simply press "M" to begin the MODIFICATION process.
MODIFICATION SCREEN

In order to create a lower cost schedule, each of 3 subgoals must be
achieved IN THE ORDER LISTED. Cost reductions grow less frequent as you
proceed through the stages. When you have finished all three stages, press "F"
for FINISHED. (Press PgDn for more detailed explanations of each stage.)

1) REDUCE COMBINED INVENTORY HOLDING AND TARDINESS COSTS OF SCHEDULE.

This step uses the "Analyze Schedule" feature as well as the
Modification Screen. INCREASES in total cost are IGNORED.

2) REDUCE CHANGEOVER COSTS BY MOVING JOBS 1 POSITION.
3) REDUCE CHANGEOVER COSTS BY MOVING JOBS 2 POSITIONS.

These steps use only the Modification Screen. Total cost should NOT
be allowed to increase. In both cases, several moves are attempted around
each changeover cost. A change is made only after seeing the effects of
ALL moves possible for a specific changeover cost. When a change is made,
earlier steps must be repeated. Duplicate moves need not be repeated.
STAGE 1: REDUCE COMBINED INVENTORY HOLDING AND TARDINESS COSTS OF SCHEDULE.

This is the first stage of modifying a schedule. Briefly, you will be moving between the Modification and Analysis screens until no negative net effects are present on the Analysis Screen. This might happen even on the first pass through the data. Each time you return from the Analysis screen, the 2 neighboring jobs with the MOST NEGATIVE net effect are exchanged. The total cost may increase as a result of one of these exchanges, but you should NOT restore the previous or best schedules even if the total cost increases more than once. Any increases will be eliminated in subsequent stages.

IF YOU NEED EVEN MORE DETAILED HELP CONCERNING THE FIRST STAGE:

After leaving help, press "A" for "Analyze Schedule." Once you are presented with the Analysis Screen, ask for Strategy Help for a more detailed explanation of Stage 1.

PgUp -- Previous Help Screen
PgDn -- Next Help Screen
End -- Leave Help
STAGE 2: REDUCE CHANGEOVER COSTS BY MOVING JOBS 1 POSITION. Screen 3 of 4

BEGIN THIS STAGE ONLY AFTER YOU HAVE COMPLETED THE FIRST STAGE (Screen 2).

The first action to take in Stage 2 is to press "B" to restore the Best schedule if the Best Achieved and current Grand Total do not match. Then repeat the following procedure for all of the changeovers from highest to lowest (some of the moves may not be possible for a specific changeover):

- Move the job ABOVE the changeover UP 1 position. Press "P" to restore PREVIOUS.
- Move the job BELOW the changeover UP 1 position. Restore PREVIOUS.
- Move the job 2 BELOW the changeover UP 1 position. Restore PREVIOUS.

IF NONE OF THE MOVES REDUCED TOTAL COST (the Best Achieved is the same as the current Grand Total), try the above procedure with the next highest changeover cost. OTHERWISE, press "B" to restore the Best schedule and then START ALL OVER AGAIN with the highest changeover cost in the revised schedule.

YOU ARE FINISHED WITH THIS STAGE when you have tried moving the jobs around every changeover (from highest to lowest) and no move reduced Total Cost. Particularly toward the lower changeovers, moves tend to be repeated. YOU NEED NOT REPEAT MOVES WHICH YOU ARE SURE YOU HAVE DONE BEFORE.

PgUp -- Previous       PgDn -- Next       End -- Leave Help
Help Screen             Help Screen
STAGE 3: REDUCE CHANGEOVER COSTS BY MOVING JOBS 2 POSITIONS.

This is the final stage of modifying a schedule. BEGIN THIS STAGE ONLY AFTER YOU HAVE COMPLETED THE SECOND STAGE (refer to the previous help screen if you have not). Reductions in cost do not occur as often at this stage, but they do sometimes occur. Stage 3 is like the previous stage, but the moves to attempt for each changeover (from highest to lowest) are:

UP: Move the job ABOVE the changeover up 2 positions (if possible).
    Move the job BELOW the changeover up 2 positions (if possible).
    Move the job 3 BELOW the changeover up 2 positions (if possible).

DOWN: Move the job 3 ABOVE the changeover down 2 positions (if possible).
       Move the job ABOVE the changeover down 2 positions (if possible).
       Move the job BELOW the changeover down 2 positions (if possible).

Again, you need not repeat moves which you are SURE you have already tried.
IF ANY OF THE MOVES REDUCED TOTAL COST, restore the Best schedule and GO BACK TO THE BEGINNING OF STAGE 2. YOU ARE FINISHED WITH THIS STAGE when you have tried moving the jobs around every changeover (from highest to lowest) and none of them resulted in a reduction in total cost. Press "F" when you are FINISHED with Stage 3.

PgUp  -- Previous
      PgDn  -- Next
      Help Screen
      End  -- Leave Help
      Help Screen
ANALYZE SCHEDULE DATA

GOAL: IDENTIFY CHANGE IN SCHEDULE WHICH WILL MOST REDUCE COMBINED INVENTORY HOLDING AND TARDINESS COSTS OF SCHEDULE.

The Analysis Screen is used ONLY in the first stage of schedule modification. None of the other screens listed on the menu need be consulted. After leaving help, begin by highlighting the bottom 2 jobs of the schedule. Press "E" for "Evaluate Net EFFECT." Move the highlight up to the next pair of jobs and again press "E." Continue this process until all of the pairs of jobs up to the top of the schedule have been evaluated. Note the "Best Computed Net Effect" listed at the right of the screen.

IF THE BEST COMPUTED NET EFFECT IS A NEGATIVE NUMBER, press "B" to highlight the "BEST Net Effect." Note which 2 jobs are highlighted. Press "M" to "Return to MODIFICATION Screen" and reverse the order of the 2 jobs. DO NOT WORRY EVEN IF THE TOTAL COST OF THE SCHEDULE INCREASES -- DO NOT UNDO THE MOVE. Once the change has been made, again press "A" to "ANALYZE Schedule" and return to the Analysis Screen. Repeat the evaluation process described above. WHEN THE BEST NET EFFECT IS NOT negative, press "M" to "Return to the MODIFICATION Screen" and go on to Stage 2 (consult Help again if necessary).

PgUp -- Previous PgDn -- Next End -- Leave Help
Help Screen Help Screen
Appendix 3.

Exit Questionnaire
Page 1 of Exit Questionnaire--Strategy Training

NAME__________________________________________________________

DID NOT               FOLLOWED               FOLLOWED
FOLLOW                ABOUT                 STRATEGY
AT ALL                HALF                  COMPLETELY

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Using the above scale, how closely do you think you followed the problem solving strategy suggested in the training session (a rating of 7 means that you solved the problem exactly as suggested in the training session):

1  2  3  4  5  6  7

2. If you did not solve the problem exactly as suggested in the training session, please explain why you did not.

3. Were there any aspects of the suggested strategy which were particularly difficult or confusing?

4. Briefly describe your strategy for solving the job sequencing problem.
Page 1 of Exit Questionnaire--Operations Training

NAME________________________________________

1. Briefly describe your strategy for solving the job sequencing problem.

Please explain any of the following aspects of your strategy which you have not described above:

A) How did you decide which job should be moved?

B) How did you decide on the job’s new location?

C) How did you decide when you were finished making changes to a schedule?
Page 2 of Exit Questionnaire--Both Treatments

Consult the sample Schedule Comparison Screen on the facing page to answer the following:

1) Did you use the Job Sequence information (the columns of letters)?
   YES [ ]    NO [ ]

2) Did you use any of the subtotal information concerning Inventory, Tardy or Changeover charges?
   YES [ ]    NO [ ]

3) Did you use the TOTAL charge information?
   YES [ ]    NO [ ]

IF YOU ANSWERED YES to any of the above questions, describe briefly how you used each type of information (please refer to each type of information by name in your response).

4) Describe any problems you encountered in trying to use the information on this screen.

5) If there was any specific information which was not presented on this screen which you wanted, please describe the information and how it could have been used.
Page 3 of Exit Questionnaire--Both Treatments

Consult the sample Modification Screen on the facing page to answer the following:

1) Did you use the information in the Changeover Charge column (excluding the total)?
   YES [ ]       NO [ ]

2) Did you use the information in the Tardy Charge column (excluding the total)?
   YES [ ]       NO [ ]

3) Did you use the information in the Inventory Holding Charge column (excluding the total)?
   YES [ ]       NO [ ]

4) Did you use the information in the Period Completed column?
   YES [ ]       NO [ ]

5) Did you use any of the subtotal information appearing under the Changeover, Tardy or Inventory columns?
   YES [ ]       NO [ ]

6) Did you use the GRAND TOTAL?
   YES [ ]       NO [ ]

7) Did you use the BEST ACHIEVED?
   YES [ ]       NO [ ]

IF YOU ANSWERED YES to any of the above questions, describe briefly how you used each type of information (please refer to each type of information by name in your response).

8) Describe any problems you encountered in trying to use the information on this screen.

9) If there was any specific information which was not presented on this screen which you wanted, please describe the information and how it could have been used.
Page 4 of Exit Questionnaire--Both Treatments

Consult the sample Analysis Screen on the facing page to answer the following:

1) Did you use the information in the Changeover Cost column?  
   YES [ ]  NO [ ]

2) Did you use the information in the Inventory Cost column?  
   YES [ ]  NO [ ]

3) Did you use the information in the Duration column?  
   YES [ ]  NO [ ]

4) Did you use the information in the Tardiness Cost column?  
   YES [ ]  NO [ ]

5) Did you use the information in the Periods Late column?  
   YES [ ]  NO [ ]

6) Did you use the I&T Net Effect?  
   YES [ ]  NO [ ]

7) Did you use the Best Computed Net Effect?  
   YES [ ]  NO [ ]

IF YOU USED EITHER TYPE OF NET EFFECT INFORMATION, please explain (conceptually) what the I&T net effect represents as well as how you used it.

IF YOU ANSWERED YES to any of the other questions above, describe briefly how you used each type of information (please refer to each type of information by name in your response).

8) Describe any problems you encountered in trying to use the information on this screen.

9) If there was any specific information which was not presented on this screen which you wanted, please describe the information and how it could have been used.
Page 5 of Exit Questionnaire--Both Treatments

IF YOU USED THE Changeover Costs Screen (a sample appears on the facing page), consult the sample to answer the following:

1) Briefly describe any problems you encountered in trying to use the information on this screen.

2) If there was any specific information which was not presented on this screen which you wanted, please describe the information and how it could have been used.
Page 6 of Exit Questionnaire—Both Treatments

If you used the High/Low Values Screen (a sample appears on the facing page), consult the sample to answer the following:

1) Did you use the Inventory Holding Cost information?
   YES [ ]    NO [ ]

2) Did you use the Tardy Cost information?
   YES [ ]    NO [ ]

3) Did you use the Slack Time information?
   YES [ ]    NO [ ]

4) Did you use the Processing Time information?
   YES [ ]    NO [ ]

5) Did you use the Changeover Cost information?
   YES [ ]    NO [ ]

6) Did you use the Smith’s Rule Ratio information?
   YES [ ]    NO [ ]

7) Did you use the Weighted Slackness Ratio information?
   YES [ ]    NO [ ]

IF YOU ANSWERED YES to any of the above questions, describe briefly how you used each type of information (please refer to each type of information by name in your response).

8) Describe any problems you encountered in trying to use the information on this screen.

9) If there was any specific information which was not presented on this screen which you wanted, please describe the information and how it could have been used.
Page 7 of Exit Questionnaire--Both Treatments

Completeness
Disagree Neutral Agree

1 2 3 4 5 6 7

Considering the SYSTEM AS A WHOLE use the above scale to rate the following statements. Please answer with respect to the data provided on the screens. Do not include menu items, instructions provided during training, or information accessed through the system’s help feature. If you feel that your rating merits further explanation, use the space provided below the statement.

A) Given the strategy I used, it would be extremely difficult to make job sequencing decisions without at least the data presented on the screens.

1 2 3 4 5 6 7

B) Given the strategy I used, I had to perform extremely complex mental calculations in order to use the data on the screens to make job sequencing decisions.

(Numbers appeared on the exit questionnaire beneath each item as in item A above.)

C) Given the strategy I used, the data presented on the screens is sufficient to make job sequencing decisions.

D) I found it difficult to determine what data on the screens was essential for making job sequencing decisions.

E) If using THIS SYSTEM to create schedules were a critical part of my job, I would continue to employ the same strategy that I have been using (assume that you are evaluated on the cost of the schedules you create and are under virtually no time pressure).

Please elaborate:
Page 8 of Exit Questionnaire--Both Treatments

Considering the SYSTEM AS A WHOLE, use the above scale to rate the following statements. Please answer with respect to the data provided on the screens. Do not include menu items, instructions provided during training, or information accessed through the system's help feature. If you feel that your rating merits further explanation, use the space provided below the statement.

A) What portion of the data presented can be used directly, without any mental calculations, to make job sequencing decisions using the strategy you developed?

   1  2  3  4  5  6  7

B) What portion of the data presented on the screens is essential to making job sequencing decisions using the strategy you developed?

   1  2  3  4  5  6  7

C) What portion of the data on the screens do you feel you understand on a conceptual level (that is, you understand what the number represents and essentially how it is calculated)?

   1  2  3  4  5  6  7
Appendix 4.

Data Excluded from Analysis
In addition to unavoidable exclusions (e.g., missing values on the exit questionnaire, incomplete job sets), the following discretionary decisions were made.

**SUBJECTS EXCLUDED FROM ALL ANALYSES:**

Four of the 46 subjects who participated in the experiment were excluded from all analyses. Two subjects in the operations training treatment were excluded because they had missed the class lecture on job scheduling. Two students in the strategy training treatment also missed the class lecture, but they were retained in the analysis because little domain knowledge is necessary to follow the instructions provided during strategy training. One foreign student in the strategy training condition was excluded from the analysis because of his expressed difficulty in understanding English. Another student in the strategy training condition was excluded from the analysis because of his expressed difficulty in seeing the computer screen and the projection screen used during training. The subject explained during the individual protocol session that he was having difficulty in seeing with his new bifocals.
SUBJECT EXCLUDED FROM TIME AND
USER LOG ANALYSES FOR JOB SET 1:

One operations-trained subject was excluded from
the analysis of time performance for the first job set
only. A program bug caused the system to abort shortly
after the subject had completed the first job set.
Because the subject may have worked from memory after
the system was restarted, the time measure and
keystroke data obtained for the first job set may have
been invalid.

SUBJECT EXCLUDED FROM ANALYSIS OF USER PERCEPTIONS:

One subject was excluded from all analyses of
questionnaire perception items because the middle
response was marked for every item (the only
questionnaire for which this was observed).

SUBJECTS EXCLUDED FROM ANALYSIS OF
OVERALL SATISFACTION WITH STRATEGY:

Three strategy-trained subjects were omitted from
the analysis because they did not respond to the
question as stated. All three indicated that they
would automate the decision. The subjects avoided the
question of how they would behave if they were using
the system.
Appendix 5.

Protocol Transcripts
KEY TO TRANSCRIPTS

<table>
<thead>
<tr>
<th>Transcript Number</th>
<th>Treatment Group</th>
<th>GEFT Score</th>
<th>Delay until Protocol Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Strategy</td>
<td>11</td>
<td>5 days</td>
</tr>
<tr>
<td>2</td>
<td>Operations</td>
<td>17</td>
<td>3 days</td>
</tr>
<tr>
<td>3</td>
<td>Strategy</td>
<td>14</td>
<td>2 days</td>
</tr>
<tr>
<td>4</td>
<td>Strategy</td>
<td>6</td>
<td>4 days</td>
</tr>
<tr>
<td>5</td>
<td>Operations</td>
<td>18</td>
<td>4 days</td>
</tr>
<tr>
<td>6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Strategy</td>
<td>9</td>
<td>2 days</td>
</tr>
<tr>
<td>7</td>
<td>Strategy</td>
<td>15</td>
<td>9 days</td>
</tr>
<tr>
<td>8</td>
<td>Operations</td>
<td>5</td>
<td>5 days</td>
</tr>
<tr>
<td>9</td>
<td>Operations</td>
<td>5</td>
<td>2 days</td>
</tr>
<tr>
<td>10</td>
<td>Operations</td>
<td>2</td>
<td>2 days</td>
</tr>
<tr>
<td>11</td>
<td>Strategy</td>
<td>1</td>
<td>3 days</td>
</tr>
<tr>
<td>12</td>
<td>Operations</td>
<td>11</td>
<td>4 days</td>
</tr>
</tbody>
</table>

<sup>a</sup>This subject was absent the day of the lecture on job scheduling.
TRANSCRIPT 1

(1) Yeah, I’m trying to remember. OK. (2) I think we go to the "M", Modify Schedule. Right. (3) Analyze the schedule. (4) And the "E" for evaluate net effect. (5) OK, and you push the "B" for the best.

(6) Now you write it down, and you go back to the -- what the heck was it?-- modification screen, the "M".

(7) Yeah, OK so now come down and move the "A" over the "G." (8) And now you go back to the analyze schedule, and you start the process all over. (9) With the "E", up, "E", up, "E", up. (10) We know what that is. (11) Oh, wait a minute--no, we don’t. OK.

(12) What I noticed on all of these was that we had one letter on the top that kept moving down.

(13) Back to "A." (cough) Excuse me. (beep) Oh, OK.

(14) Back to modify. Sorry about that. (15) Now this is where it got confused, because it went up.

(16) And I believe we go back to the "A" and we continue to do this same thing. I’m not sure if I dosed off when you were explaining this or what. (17) You just continue to repeat this until you get it right. (18) I don’t know. (19) Push finished. (20) Although it seemed-- I thought I had a lower score than that Saturday.
TRANSCRIPT 2

(1) ...the different orders and figure out which one is actually giving me the least. (2) Although the computer does it for me going to the next screen, I like to look and see where it actually is. (3) So it would be like at 30 or 6387 would be the least, the Random, Least Changeover, Weighted Slack, so Shortest Processing Time gives me the least. (4) I change over to the modify. (5) In looking at this, I wouldn't do any changes until I went into the analyze. (6) In the analyzing I go through and I analyze or evaluate the net effect for each one of them. (7) Gives me a -34, 10, C to B would be 150, D to C was positive, A to D is 120. (8) This gives me the net-- the largest change is 100, so I go back to modify and I change these 2. (9) Calculating a new-- and I anticipate that it will be reduced from the original number, which it is, by less than 100, which I figured that it would be less than 100 because of other charges, but I go back and take a look at what I’ve done. (10) Reevaluate-- let’s see, where am I at?-- analyze schedule. (11) I go back through and reevaluate. (12) I know the first one’s 30. (13) I don’t have to do it, but I’ll do it just to get the lowest one on the screen. (14) 34. (15) That shouldn't have changed. (16) That
shouldn't change. (17) This should not have changed. (18) And then the G to D will change. (19) And there’s a change there. (20) A to D should be positive 100, and it is. (21) So I go back. (22) My next one is G to D, so I can move it down and get a change it appears. (23) Let’s see, inventory cost versus tardiness is what I’m charging. (24) Job duration, changeover cost I’m not worried about yet. (25) So modify. (26) G to D. (27) Should give me a lower cost unless changeover is higher. (28) And it does. (29) I go back and do the same thing again. (30) I know that’s 30-- negative 34. (31) And I still have a changeover here so I’ll go back and do that. (32) This should be positive. (33) It is. (34) And this should be positive. (35) And it is. (36) Modify. (37) G with... (38) Eventually this will catch me and the changeover cost will be more and it will give me a positive number versus negative...like this time. (39) That was a bad move. (40) So I go back and do-- bring it back to where I was at. (41) I’ll do E to F because that also had a negative number in the analyzing. (42) I’ll change it now to see if that effects it. (43) And it in fact did effect it. (44) I’ll go back in and check and see
if this change made the effort, changing E to F. (45) So this should be positive 34, E to F. (46) Let's see, evaluate. (47) And it is. (48) B to F should change and it could be a negative, but it's not. (49) OK, the net effect is 46 so I go to the changeover table to see if the changeover cost between F to B is cheaper than B to F... and C to F and C to B. F, D, G, C, B, F, E. OK. (50) Let's see. Job B to F is 158, so B to F is 158. (51) F to B is 213 so that would be a positive change. (52) C to B is 114. (53) B to C is 127. (54) That's a positive change. (55) So that wouldn't affect it. (56) That would increase it. (57) So I don't need a change there. (58) I've got periods late. (59) Job A is done ahead of time, D is done ahead of time, G is done ahead of time, C is done ahead of time so there's no tardiness charges there. (60) B is done two weeks late so there is a charge for tardiness of 22 or 44 dollars times 2. (61) Job F is done early, job E is done 1 week late, so I'm only 3 periods late total for the whole thing. (62) The cost of inventory I would evaluate to see if there's any way that I could move the inventory levels up to reduce the amount of time they're done ahead of schedule for inventory costs. (63) So checking the best, or excuse me, evaluating net effect of changing G to C says that
I get a negative change there. (64) So I should change G and C possibly to move the inventory cost of 35 dollars, reduce the amount of time that it’s late. (65) So let me go see, that’s an 89 dollars between G and C. OK. (66) Modify. (67) I think I did this before but I don’t remember. (68) OK. I did that one before. (69) The reason it went up is the changeover cost was more costly. (70) So I go back to the best achieved. (71) Analyze again. (72) Taking a look and see if there’s any other places I can move. (73) This is 7 weeks early, 36 dollars a period. (74) This is 2 weeks late. (75) If I change B to F, I’ll check and see if that physical change will make a difference. (76) It’s easier for the computer to do it than me to do it, so I’ll let it do the work. (77) And it went up. (78) Hmmm. Let me go see why. (79) Went from 19 weeks to 16 weeks but the changeover cost must be more. (80) And I’ll check changeover cost between B and E and see if there’s a way to reduce that. (81) B and E—so B to E is 170, E to B is 267. (82) If I change that that’ll increase it by almost 100. (83) But I am 16 weeks, 1 week. (84) I don’t think that will do any good, but I’m going to try it anyway. (85) Should go up. (86) It goes up. (87) Go back to the
best achieved. (88) OK, I'm not tardy, but I'm early. (89) OK, check the order is A,D,G,C,B,F,E. (90) I want to go back and check and see what the original computer picked for the shortest time. (91) I want to see what the other ones were. (92) And that is compare schedules. (93) This is where I wish you had my choice total. (94) I could see how I compare. (95) OK, A. So I basically moved A to-- G to A. A,D,C,B,F,E. (96) I changed shortest, more tardiness, more inventory cost. Let's see. A,D,G. OK. (97) All I can say is I'm looking. (98) I have no idea what I'm doing. (99) I'm just looking at this point. (100) I'm going to play for a minute to see what would happen if I did something. (101) This is 19 weeks early. (102) I want to see what happens if I move it down to where it would be in the line of 19 weeks. (103) Job duration is 18, so that would be 11. (104) Moving past here would be-- so that would be 10. (105) So if I moved C to here I just want to see what would happen. (106) Just to see if there's a cost change...because it's way early. (107) And inventory cost is what seems to be costing me. (108) And that hurt alot. (109) I'm not late but-- OK. Analyze. What did I do wrong? (110) OK. It's almost impossible for me to do all the jobs at the period
they’re due. (111) So I have inventory cost no matter what I do. (112) The shortest processing time initially gave me the lowest cost. (113) By adjusting a couple of them down for changeover cost I’ve reduced the cost. (114) (beep) Ooops. Go back to the modify. (115) Change back to best. (116) Zero tardiness cost. (117) 144 inventory because I’m early. (118) Let me go back and do one more evaluation to see if I missed anything. (119) Evaluate net, evaluate net, evaluate net. (120) I’ve done that one before. (121) That’s a 74 dollar increase and a 68. (122) I’ve already changed those to make them positive. (123) It’s a 74 dollar change. (124) If I go to the changeover table D and G. (125) If I go from D to G it costs me 256. (126) If I go from G to D it’s 216, so that’s not the 74 dollar delta but there is an improvement. (127) If I go from A to D it is 618. (128) From A to G is 126. (129) So there’s another 10 dollars still doesn’t overcome it. (130) And the last is G to C is 162. (131) D to C is 300. (132) So that would increase it. (133) That won’t do any good. (134) Let’s see if there are any differences in High/Low Values, although I don’t really see any use of this screen. (135) A,D,A,D, they
TRANSCRIPT 3

(Accidentally recorded over first few moments of protocol)

(1) ...and I’m going to go back. (2) I see that I’ve made an improvement, and I’m going to go back to the analyze screen and continue this until I get rid of all my negatives which is also the instructions you gave me. (3) I really don’t even need to do these again because I knew I knew that was a -34, but I have to do all of them because the top 2 are going to change. (4) Might be... looks like -74. (5) OK. So I’m going to press "B" for the best which was -74 and switch those 2 jobs. (6) Going to the modification screen. (7) D is to go above G. Again, everything that I’m stating here is everything that you showed us in the exact order at this time. (8) OK. Go back to analyze. (9) Still have a negative, so I know I have another round of this. (10) Just moved another job. (11) Going back have to analyze again the schedule. (12) I still have negatives. (13) Going to modify B above G. (14) Go back again. (15) I’m pretty sure I still have negatives. Yep. (16) But I can see my numbers becoming larger into the positives. (17) OK. This should be my last one. (18) F before E. (19) But I’m going to analyze the schedule one more time to make
sure. (20) Positive, positive again, positive, positive, positive, positive. (21) OK. So I’ve now rearranged the jobs and I have I have all positive net effects and the next step is to look at changeover costs. (22) I just got done concentrating on inventory and tardiness which is basically what you said to concentrate on first. (23) And I’m looking for the highest changeover cost which is 300 and I’m going to change jobs that are around the 300. (24) I’m just going to jot this down here. (25) I’ve got 300 is the largest, D,A,C and B and I can basically make... (26) Let’s see, I can move this... I can make 3 moves here around that largest changeover cost. (27) I’ll do one at a time and go back to... Oops. Hang on a second. (28) I need to I need to do something here. (29) I need to go back to the modification screen before I proceed and I need to restore my best schedule before I continue on. (30) OK. Now I can begin moving these jobs. (31) I’m going to start by moving in small increments which is another instruction you gave us so not to make such drastic changes in the tardiness and inventory cost as I change the changeover costs. (32) OK. My previous was better but I’m still going back to my previous because I have to calculate
the other ones. (33) I’m going to move C next. (34) I did better this time, but I still have to go back to previous and do one more change in the job arrangement. (35) I’m going to take D above A. (36) OK. It looks like that one came out the best, so I’m going to restore that. (37) I don’t know if the system automatically does that so I’m just being safe here. OK. (38) My numbers now my changeover charges in that area have gone down into the 100’s. (39) I do have one large one, 249, so I’m going to attack that next. Again, per the instructions you gave. (40) And I can only make 2 changes here I believe because it’s towards the bottom. (41) Start by moving this job here. (42) I didn’t do better. (43) I got worse but I’m still going back to previous and then I’ll move the next one up. F above G. (44) And that got worse, so I’m going to get my previous up. (45) I’m going to just hit restore just in case. (46) I just want to make sure it restores that as the best schedule. (47) So there’s really I can’t get rid of the 249. (48) I can go to the next highest changeover cost which is 197 and I can make some changes here. (49) I can start by moving F above G. (50) I did not do better. (51) Go back. I can move... (52) Oops. See now I did that move twice now that I look at my sheet. (53) I
already did that. (54) G. I did not move G, though. (55) Let me try moving G. (56) I did not do better. (57) I’m going to write these down now so that I don’t do that again where I did the same calculation twice. (58) And B to C I already did so I don’t need to do that again and the next one is 136. (59) I’ve already made E. (60) I’ve already changed F. (61) Let’s see, 114. (62) I’ve already changed most of them I think. E,F,G,B,C,D,A. (63) So I’ve really completed the short moves with the changeover costs so next thing I can do is make double moves. Again, this is your instructions as well, basically, to just move them next to double moves. (64) So I’m going to go back down to the 249 which is the largest changeover cost. And, basically, everything I’m saying is restating your exact instructions which I paid attention is what I’m proving to you. You didn’t bore me totally. I did pay attention. I think its the extra credit that you mentioned that kind of kept be awake that day. (65) OK. So per your instructions I’m moving to E. I don’t feel like this is my decision process. Can I just say that? It’s yours. It makes sense. You gave me full instructions. I’m just following them. (66) There’s two moves there. (67) And I didn’t do any better.
(68) Try jumping the next one up two moves. (69) And I did not do better. (70) I'm looking at the totals which is really where I've paid most of the attention the whole time. (71) Moving the next one up two.
(72) Oops. I think I made an error here with the arrow. (73) I'll see when it comes back. (74) Yeah. I accidentally didn't skip two jobs. (75) I went one.
(76) So let me go back and do that again. (77) I should have gone above the A. (78) Nope. UhUh. Not making any improvements here to the schedule. (79) What I should have done, though, is-- I went all the way up-- I should have gone back and still stuck near the highest numbers and came down. (80) There's up movements and there's also down movements with the jobs. (81) So I'll go back down to where I should've... (82) (beep) Oops. Wait. Do I hit escape?. OK. (beep) Hey, why can't I do that? Oh, OK. The wrong arrow. The wrong way. (83) I'm back doing the double jumps but I'm going down versus up and I'm not seeing any improvement. (84) I've got another one here to do. (85) Nope. Basically I think this is it. (86) I've made single movements. (87) I've made double. (88) You didn't say to go to triple so I'm not going to, but basically I think the further you jump the less improvement that you see so I think
that's why we stopped there. Not much of my own thought process here. (89) Should I continue to another job?

**EXPERIMENTER:** If you're done with this job set, yes.

**OK.**

(90) I think I've seen this number before so I'm fairly certain that I've completed this. I'll just say 6. I'm not totally. (91) And my strategy I'll give a 6. There's room for human error.
TRANSCRIPT 4

(1) OK. Now I don't know if I cheated when I did this but I did not use this schedule here. (2) I went on to the next page. (3) Well, modify the schedule. (4) And I went directly to analyze schedule. (5) OK. Now I use the net effect and the I&T. (6) "E." (7) Do I have to talk this out even as I go here?

EXPERIMENTER: If you're thinking say what you're thinking.

(8) Oh, OK. I'm not thinking really. (9) Now I press "B" for best net effect and I move on to... (10) I want to see the best net effect for B. (11) Oh, OK. I did that. (12) Compare Schedules...Now I go to...I forgot how to do this. (13) I modify screen with the best. (14) OK. The highest number is 300 and I use the first job above it and I move the job out and I go up one and place it here. (15) And this is the best one the best achieved. (16) And, let's see, it changes my schedule so now I go back and I analyze schedule. (17) And I do that I&T. What is it E? Evaluate. (beep) Oh. E. (18) I forgot this is slow so I have to take my time with it. (19) Go with the best and now I'm going to modify... (20) This is the best I want modify. (21) The largest number is 271 so
I’m going to go down to B which is E which is one above take it out and move it up one. (22) OK. Go back to the best one. (23) The largest. OK. (24) Go one below. (25) Take it out. (26) Move it up one. (27) This is the best achieved. (28) Analyze schedule again with the new best. (29) "E." I’m hoping I’m doing this right. Best. (30) Now I have to do it again. (31) Modify. (32) Largest number is...I’m repeating myself here. (33) I’m writing down what the largest changeover charge is. (34) And I’m going to start moving jobs. (35) Taking the job above it and I’m moving it up one. (36) That’s too high so I’m going back to the best achieved. (37) I’m going to the largest number the next largest number is 216. (38) G is above. (39) I can’t move that. (40) D is under it. (41) I’m not sure if this will change it but I’m moving it above. (42) It changed it to a lower grand total so I’m going with the best for this one. (43) Now I’m going to the next highest number which is 258. (44) The job above it I’m going to take out and move it up one. (45) That was too high so I’m going back to the best. OK. (46) I’m going to the job below, taking it out and moving it up one. (47) That’s too high, so I’m going back to the best.
(48) OK the next highest number is 144 which I'm writing the numbers down. (49) I tried D and G with 256 so I'm going to take the letter the job below it and I'm going to move it up one. (50) That's lower, so I'm going to use this as my best achieved. (51) 140...Umm. I'm losing track of what I'm supposed to be doing here. I'm going to strategy help which is "S." I'm paging down to continue reading the strategy. Use the modification and analysis screens until there are no negatives-- I did that. I forgot what I did Saturday so I'm writing down the way I should move the jobs. One above, one below, two below. And I move them up one position. OK so I'm going back. Where's End? OK so I did 114... (52) I'm starting over with my moves because I think I messed up somewhere. (53) So I'm going to the highest number changeover and move my jobs. (54) The one above. (55) Going back to my best. (56) One below the highest job one above. (57) Go back to my best. (58) The next highest job is 162. (59) Writing this down so I won't forget. (60) The one above and move it one position up. (61) Go back to my best. (62) One position down. (63) Go up one. (64) Going back to my best. (65) Two positions down up one. (66) Going back to my best. (67) I'm going to my next highest number which is 158. (68) One
position above goes up one. (69) Going back to my best. (70) 158 up...moving it up one position. (71) Going back to my best. (72) OK. Two jobs down. (73) Taking it out and moving it up one position. (74) Going back to my best achieved. (75) My next highest number after 158 is 140. (76) D can’t go up any higher. (77) A is below it so I’m going to take A out and move it above D. (78) I don’t know if this is going to change it but I’m going to try it. (79) That’s higher so I’m going back to my best. (80) Two positions down is G. I’m going to take that out and move it up one. (81) That’s higher so I’m going back to my best. (82) And the next highest number is 126. (83) I did A with 140. (84) I did G with 140. (85) And I did C with 162 so my best is 6029 so I’m going to hit "F" for finished. (86) Are you completely finished making changes to the schedule? Yes. (87) Using the 7 point scale shown, how sure are you that you found the lowest cost schedule possible? I’m not very sure that I found it. I’m going to say about 4. (88) Using the 7 point scale-- I know I forgot a st-- I don’t know. Can I go back? I forgot to move positions twice for the lowest. I just remembered. How sure are you that the strategy...I know I skipped a step so I’m going to hit 2 here. Are the values...yes.
TRANSCRIPT 5
(tape recorder not recording for first few moments)
(1) I just keep going until I don’t find any
negatives and do the net effects. (2) Oops. I went
too far. (3) OK. So I don’t have any negatives left
so I’ll see what I can play with the holding and the
tardy to move them more than one. (4) E is tardy
alot. (5) I’ll try to move that back. (6) I think
this is the best I came up with last time. (7) Try
things and if they don’t work change it back. (8)
Like B. B it’s got a tardy charge but the holding
charge isn’t very big yet. (9) I’ll move that up.
(10) Or maybe not. (11) I’m going to write down the
order they’re in and make sure that I don’t have any
really high changeover. (12) This one changeover is
high from D to G but I don’t think it’s any better
anywhere else. (13) That’s better. (14) I made some
more changes. (15) I may as well go back and see if I
have any negative net effects. (16) Well, I’d better
quit here. (17) I get bored if I work with them too
long and I don’t make any progress.
TRANSCRIPT 6

Stages...

EXPERIMENTER: Try to talk out loud.

(1) OK. Well, I remember from Saturday that there are three stages I have to go through to find the lowest cost so I'm going to... (2) This has already been analyzed so... (3) I've got to play with it for a minute because I don't exact... (4) OK I'm going to analyze the costs so... (5) I'm going to evaluate the net effects. (6) (beep) Oops. Oh. Now I want to see which grouping gives me the best net effect. (7) OK that was the -100. (8) Let's see. Got to return to the... Oh. Going to return to the modification screen. (9) Change G and A. (10) I modified the schedule by switching G and A around which gave me a better cost than what I previously had. (11) Now, Stage II. I should have brought my notes because I forgot what I'm supposed to do.

EXPERIMENTER: Remember that there's strategy help there.

OK. Let's see. (beep) Oops. My problem with this program is I forget what I'm supposed to all the different steps unless I write them down and go step by step. Saturday I found out that if I wrote down Stage
One do this, this, this and that's better for me
because I don't remember. OK. Let's see. Do this
one. Another problem that I was faced with Saturday I
was not sure that I could touch the two jobs that I
originally moved in the first stage. When I got to the
second stage I didn't know that you could bother those
again. (12) Oh, now I think I've forgotten to go back
to my previous screens. (13) That blows my other... I
want the best. (14) OK. Now. OK. Ooh. That's good.
(15) Well the best I've achieved is 6029. (16) I
think that's the best. I think I've gone through all
three stages so I am going to do what? (17) Oh. That
wasn't what I wanted to do. (18) (beep) OK. Now how
do I get out of here? Yes.
TRANSCRIPT 7

(1) OK. Analyze the schedule comparison. (2) OK. Smith’s Rule is out of order. (3) Shortest Processing Time, Weighted Slackness Rule, Least Changeover Heuristic and Random Order. (4) Looks as though shortest processing time and weighted slackness rule are very close in total cost and random order is the highest, hmm. (5) Modify. OK. (6) Oh, least changeover heuristic is the lowest. No, it’s not. (7) Let’s see. Shortest Processing Time. OK. Modify schedule. (8) OK. I get this chart the shortest processing time modification screen. (9) What I’m going to do here is go down. (10) No. Wait a minute. I’m not going to do anything. (11) I’m just going to look at this and then I’m going to go to another screen, right? (12) Wait a minute. Analyze the schedule. That’s what I want to do. (13) Restore, job to be moved, new location, delete, compute new costs, restore previous, best achieved. Let’s see. What was I doing? (14) Oh, I was making changes. Help. Help. Operating Help. (15) About the information...to view additional information before making a change press "A" for analyze. Press page down for the next help screen. (16) (beep) Hmm. That’s
it. End. (experimenter turned off the numlock which was on) Oh, OK. Thanks. (17) Use the up and down arrows to highlight the job... (18) Leave help. (19) OK. "A." I want to analyze the schedule. (20) OK. Best net effect is highlighted. Hmm. (21) Evaluate. Hmm. Evaluate. Evaluate. Evaluate. Evaluate. Evaluate. (22) Now, best net effect was that one. OK. G and A. (23) Return to modification screen. (24) Let's see if we can get better. (25) Let me see. G and A. I won't forget. OK. G and A. (26) What am I doing? I'm going to make G and A G and D. (27) Let's see. Take that. Enter. (beep) Oh, no. Left arrow. (28) That one I want to put up there. Enter. Calculating costs. (29) Oh, I forgot to look at the cost. Let's see. (30) Best achieved. Previous total. Oh, that came out better. 6368. (31) Well. I did G and D instead of G and A. (32) Now, that's it. Go back to the other one. (33) Analyze the schedule again. (34) Evaluate. Evaluate. (35) Hmm. Best net effect is now D and A. (36) Modify it. (37) I'll make that A and D or D and A again. (38) Back to 6368. Hmm. (39) Back to G, A and D. (40) I'm going to analyze it again. (41) Oh, no not analyze this. Oh, wait a minute. What did I do? No, that was right. (42) OK. Analyze the schedule data. (43) OK. What am I doing?
I&T net effect. (44) Best computed net effect.
That's what I'm looking for. (45) Now...inventory.
Ooh, wow. Seems like more periods are later now so.
(46) Evaluate. Evaluate. Evaluate. (47) Back to
where I started. Hmm. (48) Modify. No wait. What
was that? (49) Analyze again. (50) Now what was the
best one? (beep) Oh, shoot. (51) I have to do it
again. But it's G and A. (52) It's G and A. I knew
it was. OK. (53) I'm going to modify G and A again.
(54) Forget all that. I'm going to make... (55) I'm
going to do it this way. (56) I can do it any way I
want to? I'm going to make F go up here. See if I get
anything lower. (57) Hey, 6368. Same thing.
(58) That wasn't anything. (59) "P" to restore the
previous. Same thing. Hmm. (60) Analyze the
(62) G and A again. OK G and A. (63) Modify. (64) G
goes here. A goes there. Calculating cost. (65)
Hey, it's lower. OK. (66) Go back. 6317. (67) Now,
evaluate. (68) Best is G and D. (69) Now, modify
it. G and D. (70) We'll put D here and G right
there. (71) OK. 6187. Great. Now, moving right
along. Analyze. (72) OK. The best. G and C. OK.
(73) Modify G and C. (74) Take C and move it above
G. Enter. (75) Now they're just about all...Nope, they're not all negative periods late. (76) Let me see. Duration of the job: 4, 5, 8, 3, 9, 10, 10. (77) This isn't telling me anything. Let's see. (78) Evaluate again. (79) Best. (80) Figured. 71. G and B. It's going right down the line. (81) OK. Modify that. G and B. (82) And take B and move it up above G. (83) Enter and I bet it will be lower yet. Let's see. Has to be. (84) Yep. 6037. I knew it. (85) Go back to the first screen. (86) Analyze the schedule. (87) And the next one's are going to be I bet G and E. Watch. (88) Evaluate. Evaluate. Did I have that? OK. Evaluate. Evaluate. Da dah. Da dah. (89) Best. (90) No. E and F. Hmm. E and F. OK. Modify E and F. (91) We'll put F above E. Calculating it. (92) It'll be lower maybe. Maybe. (93) Yeah. 6016. Wow. How about that. OK. Go back to analyze for the last time. (94) Now either G and F or F and E again. No, let's see. Hmm. (95) Evaluate. Evaluate. None of them. (96) Hmm. Got 'em all. That's it. OK. Now. (97) We've got the best. Hmm. Forget this schedule. (98) Let's see what it looks like. (99) Everything's got negative late periods except E and that's about it. (100) Everything's negative. OK. Now. Hmm. (101) "M." Modification
screen. (102) I’m ready to do something. What am I doing? I’m doing stage two. I’m going to move two jobs and make them better right? I think that’s the next step. (103) Best achieved is 6016. Great. OK. I’m going to take...Hmmm. (104) Oh, wait a minute. Am I going to move two jobs or am I going to move one job two spaces? I’m going to move... (105) Let me see this thing now. Tardy charge. Hmm. Nothing. Oh, because there wasn’t any tardiness. Periods completed. (106) What is this? The shortest processing time modification screen. (107) OK. What I’m going to do is rearrange these and compute a new cost for the shortest processing time. (108) Oh, oh, I see. That last screen we must have got the shortest processing time. (109) This screen we’re going to fix it even better. OK. Now. Hmm. (110) Best achieved 6016. That’s what I’ve got to beat. (111) So we’ll go...Hmmm. A,D,C,B,G,F,E. (112) What am I doing? C take B. Wait a minute. (113) A, D and C fit together. If we take A...No no no. (114) Take C and move it above A (beep). (115) Oh forget this. Help help help. Let’s see. (116) Next help. Use the up and down arrows, use the left arrow, use the up and down, use the... (117) Next help. I’ve done all that.
(118) To undo a change, to indicate that you have finished, press "P," to undo a series of changes, to indicate that...No, next help. (beep) (119) No next help? Hmm. Previous help. Page up. (120) To make a change, modification screen, along with the total cost, about the information on the screen...OK, wait a minute. (121) Next help screen, along with the total cost information for the schedule currently displayed on the screen, you are also provided with total cost information for the previous schedule, this will only exist if you have made at least one change, any number of, best achieved, this number represents the lowest total cost schedule you have achieved in the course of modifying the schedule displayed on your screen, OK, press "A" for analyze. Oh, (beep) no previous? End. (122) Alright. What am I doing? I’m going to take two at a time. (123) No, I’m going to skip two. (124) I’m going to take B and move it above...D? (125) Let’s see. No no. ’I want to start with A. (126) Delete delete delete delete. Delete you too. Where are you? Uh oh. (127) OK. Start with A. I’m going to take A and move it down below C. (128) Alright. That’s taking two at a time. (129) Oh, god. Made it worse. OK, nope. Don’t want to do that. (130) Put A back. (131) No, restore. How do you do that?
Let's see. Restore the best previous. (132) I'm going to take D and move it down below B. There. (133) Nope. Doesn't work. Restore. (134) Let's see. Now I'll take C and move it down below G. (135) Shortest processing time. Nope. Inventory holding charge, tardy charge, period completed. (136) Why would I want to do that? What's this got to do with it? Operating help--Strategy help, strategy help. I need some stra--Wait a minute. (137) I did A, I moved D below B, I moved C below G. (138) Let me finish. B below F and then I'll get some strategy after I do the last one. (139) Nah, that doesn't work. Nothing works. F below...Let's see I did A below C, D below B, C below G and I did B below F and I'm going to take G and go below E. Da dah. (140) Let's see. Two at a time. (141) Oh, god. Restore. It's getting worse by the minute. This is not improving at all. Strategy help. Let's see if this will help me. Oh. In order to create a lower cost schedule, each of three subgoals must be achieved in the order listed. Cost reductions grow less frequent as you proceed through the stages. When you have finished all three stages press "F" for finished. Of course. Reduce combined inventory holding and
tardiness costs of schedule. This step uses the analyze schedule feature as well as the modification screen. I did that. Increases in total cost are ignored. I did that, didn’t I do that? OK. Reduce changeover costs by moving jobs one position? I did that. I think I did that too. Now I’m trying to move them two positions. OK. These steps use only the modification screen. Total cost should not be allowed to increase. Right. In both cases, several moves are attempted around each changeover cost. A change is made only after seeing the effect of all moves possible for a specific changeover cost. Changeover cost. When a change is made, earlier steps must be repeated. Duplicate moves need not be repeated. Now, wait a minute. Do I have to go back to that one screen? This is the first stage of modifying a schedule. You will be moving between the modification and analysis screens until no negative...OK. That happened. This might happen even on the first pass through the data. No, it didn’t happen on the first pass. Each time you return from the analysis screen, the two neighboring jobs with the most negative net effect are exchanged. Right. The total cost may increase as the result of one of these exchanges but you should not restore the previous or best schedules even if the total cost increases more
than once. Any increases will be eliminated in subsequent stages. What do I do? If you need even more detailed help concerning... Nope, nope. After leaving help press "A" for analyze. Once you are presented with ask for strategy help-- I'm through with stage one. End. Leave help. (142) OK. 6016. That was the best. OK. (143) I moved them one step and I moved them two steps and they don't get any better so I am finished. That's it. (144) You will not be able to return. Nope. Are you completely finished making changes to this schedule. Yes. (145) Wow. Oh, this thing. OK. Absolutely sure. Using the seven point scale shown above, how sure are you that you found the lowest cost schedule possible for the job set you just completed. Please respond with respect to schedules which can be created after attempting 100 or fewer changes. How sure... 7. (146) How sure are you that the strategy you used was the best way to solve the problem... Six. There had to be a better way. That was about the best. That was about the best you could do. Are the values you entered above correct? Yep.
TRANSCRIPT 8

(1) Now I'm just kind of basically looking over the data. (2) And what I did before I was looking— I went to the totals first and I just really kind of looked at which method the shortest which one was the the was the least cost and I noticed that the shortest processing time would be the least cost so then what I did was then I would switch to modify schedule and I remember OK the shortest processing time I remember from our teacher when he was telling us various various variables I remember that when you're using the inventory variable one of the one of the methods that you can use was the shortest processing time. (3) So what I did was I tried to hold one variable constant and change and you know make changes within maybe this variable. (4) And so what I did was I tried to put them in order with the with the with the least cost at the beginning and then I put it in an ascending order. (5) So what I did after that, I looked at the data and then I noticed that I would try to move this 435 behind 336. (6) So what I did was brought this I would bring this down and...is that where I want to put it? Wait a minute. Wait a minute. (7) Is that what I want to do? No. No no no no. No that's not what. I wanted
to take... No. No no no no no. And then I hit I hit delete because I messed up. (beep) (8) Wait a minute. Press any key to continue. OK. That's what I needed to do. (9) I came down to this 435 and then I wanted to take that out and insert it two up. (10) OK, that's what I did and then I put it in there and then I hit enter. OK. (11) And then I looked and I saw that that total the total was higher than that so I said well, that didn't work. (12) And then what I did is I just tried it again. (13) I would go down and I just tried it again to see what would happen. (14) I pulled this 315 out and I put it here. (15) OK. OK. Well, OK. Then I noticed that my numbers kept getting bigger so I said well, I don't want to try that anymore. (16) Maybe I would try another variable and try to do that in in order. (17) And the reason why I was just trying to hold one variable const-- all the other variables constant and manipulate one particular variable was because I was just thinking in terms of in companies certain companies have certain priorities. (18) Maybe, for instance, one company may be interested in the tardy charge. (19) That may be a very important factor to that particular company and some companies want to particularly make sure that they hold down the inventory holding charges. (20) Although
they’re concerned about everything else, that may be a priority for them. (21) So I was thinking in terms of what the company what the company’s priority was. (22) So then what I did was I went to 58 and I pulled 58 out and then I put it in... so I would try to put it in order to see how that would work. (23) OK. No, I know that’s not that’s not going to work because that’s much much larger. (24) Then I moved over for instance maybe to the changeover charge and I tried to work it that way and I would try to see what would happen if I took the 144...that’s not what I want to do. Can’t do that. OK. If I took the 144 out and then brought it down and put it in here. (25) I tried what I did was I tried you know to do a couple of them that way. (26) That was a little bit better than previously but that’s still way off way off base. (27) So sometimes I would go back to sometimes every now and then I would go back to best achieved answer and then start all over again but I would I looked and then I said OK well, let me try it again this way see if I could. (28) And then I’d take the 101 out and I would move that out front over the top. (29) I’d do those a few times, not necessarily, maybe about two or three of them I’d do that way. (30) Then OK I noticed that that was even
worse. (31) Then what I tried after that after I tried the then I started trying to bring the highest numbers up and I did the repeated the same process only starting off with the highest numbers. (32) So I might have taken this variable and so I said that I was going to bring this one and since that was the highest and then I tried to see what would happen if I put it at the very top. (33) Oh, OK. Enter. That’s even worse. (34) Basically, I kept doing it that was just my method I kept doing it that way until in some cases in the first two sets of the actual set I got a lower number, a slightly lower number than the best achieved answer that was that came up originally on the screen. (35) On I think the last two sets, because I did four, on the last two sets I came up with the I entered for my finished answer the best answer that I had achieved even though it was a little bit higher than the best achieved answer that originally came up on the screen partly because I had fiddled with it and fiddled with it and then it got to the point where it was just kind of after that after I saw that this the method didn’t kind of really work in all cases I kind of just got tired of it and then I just moved on to the next one. (36) And that’s basically the way I did it. (37) Do I did you want me to keep going?
EXPERIMENTER: Yeah, I want you to just finish this. There’s just one more set. Just finish this.

(38) OK. OK. Let me see. OK, then I would probably what I would probably do is go back to my previous answer. (39) And then I would just...Let’s see I had...I would probably... (40) Oh, then I would, let me see, work on the tardy charges and bring and bring that up to the top and to see what would happen. (41) OK. Then I noticed OK then I noticed that that was alot better so when I noticed it was alot better then I said well what if I try to bring one of the the the the tardy charges the the higher cost up to the top because I was thinking maybe the company didn’t care that that they didn’t really care about that. (42) So then I would probably... OK and I would take this out and I would move it I would try to move that up to the top (43) OK way up to the top so that the and we’ll see what happen would happen there. (44) OK. Well, I’ll go back to my previous total since that was better. And I would see... (45) OK then I would probably move this I would still bring it up to the top but maybe I would put it maybe in the second spot and see what would happen. (46) OK. Then I’d go back to the previous and I’d try to do the same thing bring the 123 down on the 152. (47) Basically now I’m just
going on trial and error. (48) No, that’s not going to work. (49) So I’d go back to my previous total and then I’d try maybe another leave that intact and then I would mean I would leave this column total intact and then I would probably try changeover charge. (50) No particular reason why I chose that I just. (51) OK this is just trial and error here. (52) (beep) Ooh, uhoh. OK. I’d try to take this and bring this up to up at the top and... put this here. (53) See what would happen if the changeover charges would cost more. (54) Hmm. Well, OK. Let me try to... (55) Well, when I did it Saturday by now I would get I’m going to get tired and just have to go with go with this previous total of 6561 as my finished answer. (56) What I did what I did Saturday I think I think I got a lower answer than this. (57) It was as low as that somewhere but I forgot to write it down now and so now I can’t go back to that. (58) Saturday I kept up with my numbers a little bit better as to what what the lowest total I had. (59) So I would just have entered finished making changes. OK. "F", "F", "F." (60) What it say...Let me go back and do this again. I should have said yes. OK. (61) OK. It says using the seven point scale above how sure are you that you found
the lowest cost schedule possible for the job set you just completed. And so I would I would say I said one. I would say one because I don’t want to say I was totally uncertain because there may be a possibility that of all of the that may be the I don’t know that... Well, we know the lowest answer was the 6300 answer but when I got in cases last Saturday when I got numbers that were lower than the best achieved answers that originally came up on the screen I still marked one because I felt as though I got a lower answer but I felt as though there might have been something even lower out there. So I just...I didn’t want to say totally uncertain so I would I just in all the cases I just chose one. (62) And this says using the same seven point scale how sure are you that your strategy used was the best way to solve the problem if the objective is to create the lowest cost schedule possible while attempting 100 or fewer changes and I said six and the reason why I said six was because and that that’s closer to the total certainty in the scale because I said six because it was really I felt with what I had to work with with what the tools and what the materials I had to work with I didn’t know I didn’t know for instance things like what the company I didn’t have other factors to use I didn’t know what the
company felt was important, I didn’t know what if you were trying to supply materials or supply things to people, I didn’t know what the suppliers wanted. So I felt that, considering that what I had to work with, that this was the best way trial and error after a while was the best way about the best way that you can do it when you don’t have alot you know a small amount of information or you don’t have all the information that you need to make a decision. And so then after that I’ll put down yes and then I move on to the next one.
TRANSCRIPT 9

(1) Modify schedule. (2) Something I didn't try before I didn't compare the jobs. (3) I'm trying to remember what I did here. (4) (?)...Let me go to help. (5) High and low values that's what I want. (6) I still need to find where the high/low values are how to compare the high and low values.

EXPERIMENTER: Did you do that Saturday?

(7) No, I didn't. I didn't compare two of them together. (8) I changed I went from A to I put A where D is. (9) I need to find out how to do the high and low values. (10) I still don't know how to do that. (11) Compare the two. (12) Part of the reason I didn't do it Saturday because I didn't see it. (13) You went over it but I didn't see it when I was doing it. (14) Schedule comparison... (15) OK. I didn't see it before. That's the reason. (16) It's there but I just didn't see it. (17) "B." (beep) (18) How do you compare the net effects? (19) This is different. It's going to take me time to figure out what I'm doing. (20) 10... B F is the lowest. (21) Is that better? -74? Is -74 higher? (22) Compare two... (23) Minus means closer to the... is minus closer to the actual?
EXPERIMENTER: Minus just means that the change in inventory holding and tardiness cost if the two jobs were exchanged is a decrease.

(24) So it's better.

EXPERIMENTER: A negative is a decrease in cost. A positive is an increase.

(25) So it's better. (26) This is higher. Let's see. (27) This is probably the best way to figure out but it takes longer to to go through it. (28) It's more quantitative to do it this way but I'm not really sure what I'm looking at here. (29) That's going down I guess I'll put this at the bottom. (30) There. 63 62 Same as it was before. (31) Oh, it's higher. Hmm. 6317. (32) So maybe I should move C. (33) It also said C was...(beep) (34) Move C next to there and see what happens. (35) Uhuh. Nope. Feel like I'm guessing now more than I am...I've been trying to quantify it but at this point I'm guessing. (36) Go back to my best. (37) I'll start plugging in numbers. (38) C. F. Go back the way I did it Saturday. (39) A... Shortest processing time this looks like the shortest processing time. (40) Try this one. See what happens here. (41) 6164. G,D,E,C,B,F... There's quite a bit of difference there. G,D is the same. E. (42) Let me try C. Try C. C,D...here. (43) Nope.
A,G,D,E,B,C,F. A,G,D's OK to keep there. (44) E let's see where E makes a difference. (45) Move it down. (46) Nope. I'll stick with what I have. (47) "B"...And what time frame? Through all of them like it was before? Oh, just two? So I can take longer with these? OK. OK. Because before I was trying to get through more but I still wanted to take enough time to... OK. I'll take a little longer then. (48) A,G,D,C,B,E...C,A,F. 6164. (49) Doesn't look like that many combinations but it's alot when you start trying to figure out what's going on. (50) 62...what did I change? A,G,D...C,F. I changed B. (51) OK. I changed F to C. (52) This is easier for me to go to the best achieved to try to get to what I want to do faster. (53) F to C didn't work. Hmmm. A,G,D,E,F,C,B. F,C,B. (54) I tried B last and it didn't do any good. B, no C. Try B in there. (55) Probably going to be higher. I just want to see. (56) Ummm. Significantly higher. OK. (57) I want to get under 6000. How do I do that? Hmmm. Hmm. (58) OK. F...-77. 19. -30. Let's see. 196. That's real high. I'm sure we can do better. 74. 100. 196. 196. (59) -30. -77. Hmm. Let's look again. 100. (60) OK, go back to this screen here. OK. A to E and I'll change
this. (61) 6740. That's no good. (62) I think that's the best I can do with this one. (63) I guess I'm through with this one. Hmm.
(1) OK. Under Smith's rules they're already in sequence. (2) So I first thing I'm going to write down the sequence already. (3) A,D,C...Oh, wait a minute. Back up. I forgot what I was doing. (4) OK. The random order sequence, is that what this is called? That gives me the lowest cost? Is that right? Naa. Hold on. (5) I'm going to go to the modify schedule. OK. (6) OK. Oh, I forgot what I did Saturday. That was a long time ago. (7) How do I get back to that other screen? (8) Delete, complete new costs, restore..."B." (9) OK. The first thing I did was try to get the best achieved schedule and I'm going to write that the jobs down in that sequence. (10) Let's see if I can work out a system again. (11) It's G,A,D,C,B,E,F. OK. (12) And the best achieved rate is 6387. (13) Now, I'm going to go to the-- what's that stupid screen? I'm going to analyze the schedule. (14) So I'm going to press "A." (15) OK. I have the schedules where I have the changeover costs, the inventory costs and all that. (16) The net effects I didn't use that last week that much so I want to go to...the high/low values. (17) OK. So what I'm looking at for the high low I want to see where the
lowest inventory holding costs were and what else did I look at? (18) A. Am I going to get nervous doing this here? (19) And then I’m also looking at the lowest tardy cost and somehow... (20) The lowest inventory cost is E, A, F and then I’m going to put B, G, and C and I’m going to go back and press a button and mess around and see if I can come up with the lowest cost. (21) OK. So I’m going to put E as the first job. (22) Can I do that on this screen? OK. How did I do that? (23) So I’m going to see if the cost for E being the lowest cost. (24) Press here? OK. (25) Oh, the cost is higher using E as the main as the first job scheduled. (26) So now I’ll go down to... (27) Well, let’s see what A will give me. I’ll put A first over E. (28) Hmm. OK. That’s a little lower. I get 6532. (29) I’m going to put G which was the original best achieved first. (30) I’m going to put that first and leave A second and E and see what I come up with this time. (31) Oh, hell. This is higher. So that blows that theory out of the water. (32) OK. Well, I’ll put C first over G. (33) Let’s just bring C all the way to the top the first job scheduled and see what happens. (34) Hmm, was too high. Hmm. (35) Well, I’m going to go back to I want to see the schedules again. (36) Hmm. Let’s see what the net effect is
with B and F which is the lower the last two. (37)
(bEEP) Oh, hell. How do I do this? Oh, a "B."
(bEEP) (38) You have not computed anything. Well,
what do I do now? (39) I want to look at the
changeover tables? (40) Hell, I don’t know. Can I do
that? (41) They all look like the changeover tables.
Let’s see. (42) OK. Let’s look at some lower numbers
here. (43) So this is just guesstimating. (44) I
don’t know what I’m going to do. (45) I’m going to
try it with, let’s see, no logic at all here. (46)
Hmm, hmm, hmm. I’m going to put F first and I’m
writing down the numbers and this is just
guesstimating. (47) F, A, B... E, and I’ll just see
how they fall. (48) I’ll just start with F and see
what happens here. Am I timed with this too? (49) F.
So I’m taking F and I’m moving it up to the top as the
first job scheduled. (50) Oh, wow. I’m getting worse.
(51) I’m going to move G back up. (52) This is
depressing. Hmm. I’m going to take A and put it back.
(53) I guess that shoots all my theories to hell.
(54) Well, it’s a little lower, but not much. (55)
Hmm. Well, let me look at this B down here. (56)
It’s the best completed period and it has a low tardy.
Let’s see what happens to that. (57) I’m going to
move that up to the top and see what I come up with.
(58) Oh, I quit. Hmm. That's higher also. (59)
Well, I'm going to go back and start with G which since
that was the best achieved so I'm going to delete G. G.
(60) Hmm, hmm, hmm. Put G first and put F second as
the second job scheduled. (61) Don't ask me why.
(62) This is depresssing. OK. We're going to just
start with the best. (63) I'm going to restore the
best achieved schedule. (64) OK, now, how come I
can't do this? (65) Hmm. F. B up there. Hmm, hmm,
hmm. (66) Let me look at these schedules again. I
don't know what for but I... (67) I like the
changeover tables. (68) So that's A follows that one.
Hmm. (69) I tried that, didn't I? Let me try another
one while I'm here. G, F-- I think I tried that
already, though. C... Let me see. (70) I'm going to
try F as the second job scheduled this time. See what
costs I come up with. (71) Well, at least it's 69.
I'm going to put C as the third job scheduled after F.
(72) That blows me out of the water here. (73) I'm
going to take C out and move it on back somewhere. I'm
going to put C last. (74) Well that didn't help it a
little bit. I quit. Hmm, hmmmm. (75) B...Well, let
me try D. I haven't done anything with D here. (76)
I'll put D after F. Hmm, hmm, hmm. Hmm. 70000. Hmm.
(77) I'm going to try B as the first one job. (78) Getting B out of there is higher. Oh, is that right?
(79) I forget where B was. I'll put it at the bottom.
(80) Can I do that? Didn't work right. (81) OK. So we got 70000 back. Hmm, hmm, hmm. Well...D. Hmm.
(82) I'm looking at the tardy time... (83) Ah, wait a minute. Delete. I'm going to make C after D job D.
(84) It went higher. Hmm. Since that's higher. I don't know. (85) I'm going to take A job A and make that last. (86) Hmm. I'm getting worse. Let me start over I guess. (87) I'll go back to the best achieved. Great. (88) Let me look at some and let me look at some of these stupid schedules again. (89) I'm going to compare schedules. (90) OK. Now let's see here. The shortest time is that, the smith's rule... A is should be up here somewhere and D... (91) 6388. OK. Let's see. Let me look at the high/lows again. (92) The lowest inventory is B and is G. OK. So it's G. Here we go again. (93) G, B I'll go with the lowest. G and then B. OK. (94) Then I'll put F, A and E.
(95) Let's see if this will work. (96) Let's go back and play around again. (97) I'm still leaving G first and I'm going to put B second. (98) Oh, my gosh. It's 75. Well that blows that out of the water. Hmm. (99)
I'm still going to put F third. (100) I'll put F second. Changed my mind. (101) F is too high so I'll have to remove F back down to the lower end. (102) I'll put F after G. (103) Hmm. OK. Let's I'll take the B out and put B-- oh, me--after A. I'm going to have G, A, B. (104) Well, that blows that theory to heck. Hmm. Well... (105) Let's take G and put that after...Put E after G. Sorry about that. (106) Man, I'm getting worse instead of better. Hmm. Hmm. (107) I'm going to go back to the best achieved again. (108) G, A... Hmm. G, A... (109) I'm going to put B before C and see what happens. (110) Did I have that before? I don't know. (111) Hmm. OK. Still high. 65. So what in the world... (112) Put B after G. I'll just do that. (113) That's worse. We'll go back to the best achieved. (114) I don't know where to go from here. (115) Let me go into those schedules again, for what I don't know. (116) OK. Let's see the net effect... (117) Let's see. Can we do it this way? (beep) Why can't I get anything? Net effect. (118) How come that won't do it? (beep) I can't get any net effects. Huh? (experimenter intervention) (119) Return. Ooh, excuse me. "E" that's why. OK. (120) A and D is 120. (121) So let's see what the net effect of G and A...How do you? "E." (122) It's a
-100 for A for G G and A. (123) OK. Maybe we're on to something here. (124) A D was 120. Let's see what D C is. (125) D C is 49 and a -100. OK. (126) And let's see what C and D. I mean C and B. Oh, that's 100 151 and the best is a -100. OK. (127) One, two, three, four. Alright. B and E is 10 and a -100. (128) And then E F. Hmm. -34 and a -100. OK. F. (129) So I'm going to put all of the negative 100's minus 100's together and see what happens. (130) That would be G, A, C, E, F. Then I had a -34. (131) Let me try that and see what happens here. So we're going to go back to the modify screen. (132) So we're going to have G, A and I want C third. (133) Hmm. 63. 63. So that's a little lower. Praise the lord. OK. (134) I'm going to put the other negatives. So one was the E so I'm going to put the E before the after the C. (135) Hmm. That blows that theory. That's higher. So we have to take the E out. (136) I don't even know where it was. (137) I'm going to put it after the B. I don't know where it was. (138) Hmm. OK. Then let's move the F. That had a -100 too I think. (139) So C I'm going to put that after the C. (140) Oh, man. That's worse too. So I have to take those F's out and I'm going to move that back down to
the end of the jobs. Hmm. B. (141) Let me see what I think B had a negative. I’m not sure now. (142) I’m going to put that after C. (143) Well, that’s higher. A 68. So I don’t want that B up there. (144) I’m going to put the B after the D. (145) OK. Now. I’m going to go back to those modifies. See if I’m on a roll here. (146) Let’s see if we can come up with something again. (147) Oh, I can start at the bottom. I want to see the net effect of E and F. -34. (148) That’s what I had before. (149) OK. B and E. (150) So it’s 10 and the best would be a -4 so that’s pretty good. (151) D and B. Mmm. -34. Mmm. (152) C and D -49 both. Hmm. (153) A and C. Hmm. Let’s see. (154) One more pair. G and A -100. (155) OK. I’m going to leave that. (156) I’m going to go back and look at this again. (157) A and C... negative—best computed net effect I&T net effect. (158) Hell, I don’t even know what that means. C and D. (159) I’m thinking about putting the D in front of the C. I don’t know. Hmm. D and B. (160) And the best is a -100. Hmm. OK. (161) Let’s go back over here to the modify schedule. (162) I’m still going to have G A. (163) Let’s try a B up here after A. (164) Oh, no. That didn’t work so I’m going to take that B back out and put it after the D. (165) A, B, C. A, B, C. I’m
sure I’ve had that D up there before, but let me try the D again after the C.  (166) (beep) Oops. The location you have indicated is the job currently... Sure is.  (167) I’m going to put it after the A. Excuse me.  (168) Hmm. That’s 6387 so that’s what it was originally, I guess. Hmm.  (169) I should be able to beat this some way. OK.  (170) We’re going to go back to the best achieved.  (171) Hmm. Have we have I had A I’m putting A over G.  (172) Ooh. That’s 6290. So that’s better and the only reason I did that was that it had a negative effect also when I did that comparison or whatever that thing was called.  (173) OK. So we’ve got A, G. Now what would happen if I put the D before the C?  (174) Hmm. Not bad but it’s not as low as... 6317. So...  (175) I want B I’m going to take the B and move that out to after G. (tape turned over)  
(176) ...and my best achieved, but I did build up my confidence because I got some lower numbers than I originally started with so...  (177) OK. A, G. Hmm. So A G had negative, so C had negative F.  (178) Well the F is at the end so let me see if I move the F up to after the C what happens.  (179) 69. So F can’t be up there. Hmm.  (180) So I’m going to move F back down
below the B but E will be the last job and see what happens. (181) Hmm. Lower yet. Was that lower? Yeah. 6222. So that’s lower. Hmm. I’m on a roll here. Let’s see. (182) So well, let’s go back over here and analyze these net effects. (183) OK. I want to start at the top and see what happens. (184) I want to see the net effect of A and G which is 100. (185) That’s weird. Before it was a -100. Oh, well. (186) G and C 89. OK. (187) I’ve got 100 and I have a -89. Hmm. Hmm. OK. C and D. (188) Hmm. There’s a -89 there. Hmm. (189) D and B. -89. (190) And then B and F. UhHmmm. 46 89. (191) And then F and E. 89. (192) OK. Now we can go back to these modify schedules again and... What happens if... (193) This is dumb. I’m going to put the C as the first job and see what happens. (194) But I want C and D. That’s 65. (195) So let’s go to D and I want that to be the second job. (196) 64. OK. Then I want B. (197) I’m trying I’m thinking now I’m putting I’m trying to put all of the negatives negative net effects together. (198) 69. But I think...G was a -89 so I’m going to put G over C. (199) Oh, hell. That blows that theory out of the water. Hmm, hmm. (200) Take G back out and put it under B. (201) Hmm. I’m going to put F as the fourth one just to see. (202) Getting worse, worse, worse,
worse. Well, we'll take A and put A at the top. Smooth it out. (203) Hmm. 6395. Hmm. A A C...
Where's my G? C D...Let's see what happens if we put G second. (204) Nah. Oh, is that the same thing I had before? OK. Hmm. Well, moving right along 6222. (205) We'll go to analyze the schedules again. (206) I'll go to high/low and see what's over there I can work out. (207) Hmm. Let's see. The lowest inventory cost A G. (208) I don't see any relationships here at all. A G. (209) A G has the lowest processing time but hey... I don't know. (210) This won't help me any more. (211) What's the compare schedules. Let's see what that is. (212) I don't know what I'm looking at here. OK. (213) Let's see what the net effects are. OK. (214) I've got 34 for F and G. (215) B and F is 34. OK. (216) D and B is 34 is the best computed net effect 34 I&T net effect C and D. (217) Hmm. C and D are negatives. That's a -49. (218) G and C are negatives. OK. 89. (219) So all my negatives are together here. (220) A G is 89. Hmm. (221) Let's look at this D B again. (222) What's that? 89. Wait a minute. C D. (223) Let's see what that one is. There's an 89. D B. 89 net effect. (224) B F 89 and F E 89. Hmm. Hmm, hmm.
(225) The only other thing I’ll try I’ll go with the lowest changeover cost and see where I end up.

(226) Let’s go with the duration of jobs the lowest to the highest. (227) I’ll put G A-- Is that the same one they had? G A...G A... Oh, that’s the way the first one was. (228) G, A, D, C, B... Oh, no, I won’t do that because I know what the cost would be. 6387. So that won’t help me any. Hmm, hmm, hmm.

(229) Let’s look at the tardiness cost per period tardy or something. OK. Let’s see. (230) B-- oh, no, C, B, A, (I don’t have my glasses on) F. Woah, E, G. Is that right? (231) So let’s go back to the modify screen and we’re going to put C as the first job. (232) Hmmm. That’s too high. That’s 6516.

(233) But let’s just go with this little theory here and see what happens. B as the second job. (234) Hmm. That’s even worse. Well, we’re going to put A back on top but I’ll still leave C and B as the second and third. (235) That’s 68. I’m going from sugar to...

Well, that theory didn’t work. Hmm, hmm, hmm. (236) Let’s just put E on as the first job for the heck of it. (237) Hmmm. That’s even worse. 79. Put A back first. (238) Hmm. That’s just as bad. We’ll go to the best achieved and start over. (239) How much more of this do I have to take? (240) I don’t think I can
do any better. 6222 is pretty good. (241) Oh, I do that too? You will not be able to return...OK, Yes. (242) Hey, I don’t know, a three. The best of my knowledge.
TRANSCRIPT 11

(1) Press "M." (2) "A" for Analyze. (3) (beep)
And it's beeping already. Come on think. (4) (beep)
It's shift. Unshift. There it is.

EXPERIMENTER: It was locked?

(5) Yeah. Let's see. You want to look for all the negatives. (6) OK. The highest net effect is G A so
I want to go to the modify screen so I can switch G to A. (7) The computer is slow. (8) And I want to go
back to analyze to see the results. (9) (beep) "E"
for evaluate. (10) Go to the best net and it's going
to be G G to D so I'll switch those. (11) I'll go
back to modify. (12) Analyze. (13) Evaluate.
Evaluate. Evaluate. Evaluate. (14) "B" for...I'll
switch C and G. (15) Go back to modify. (16)
Analyze. (17) "B" for best. (18) I'll switch G and
B. Back to modify. (19) Analyze. (20) "E." (21)
"B" for best. Switch E and F. (22) (beep). Oh,
dingaling, you don't want that. OK, you want to you've
already press any key to continue and I want to cancel
that so...move, compute new costs, restore, hmmm, move
to a new location...I should have a delete key. (23)
E and F. E and G. (25) Modify. Move E and G. Enter. (26) Hmm. My best achieved was 60...Well, I’ll ignore that. (27) Go back to analyze. (28) Evaluate. (29) This should be my last negative. (30) Move E and F. Modify. (31) Hmm. That’s even better. (32) OK, let’s go back to analyze and check that I’ve moved all the negatives. (33) Then I’ll go down and just double check that I didn’t miss anything. (34) Then I’ll go back to modify. (35) OK, Now I can move to the next step of the changeover charges. (36) I’ll start with the largest which is from D to C which is 300 and I’ll switch those. (37) Hmm. That’s better. 5992. (38) Let’s see if I can get rid of the changeover from D to B which is 271. (39) That’s not good. I don’t like that one. I’ll go back to previous. (40) Let’s see if I switch... Let’s try the E switching F and E since that’s my lar-- next largest changeover. (41) No, I’ll go back to previous. (42) Let’s go let’s switch F and G. (beep) Oops, wrong button. (43) Nope, don’t like that. Let’s go back to previous. (44) Let’s see if I do F two places. (45) I don’t like that one. Go back to previous. (46) See if I go to the next the third largest 197 which is the changeover between B and G what happens there. (47) Don’t like that one either. Let’s go back to previous. (48) Let’s go
back to the change between D and B and see what happens
if I move B to C.  (49) No, I don't like that. Let's
see if I move F...G.  (50) I don't like that total
cost. Let's move A.  (51) No, I don't like that one.
(52) I feel I have the best cost I have the best
achieved so can I move out?  (53) Have you completely
finished making changes to the schedule...yes.  (54)
Using the scale above...the lowest cost possible for
the job set you just completed. Please respond with
respect to schedules... I'll say seven.
TRANSCRIPT 12

(1) OK. Alright so the Smith's rule is the lowest.
(2) The weighted slackness is the lowest going by number 6388. 6387. (3) I'm sorry. Shortest processing time. OK. (4) Do an analyze. (5) Start looking for net effects under zero. (6) Evaluate. E F 34. (7) B E is 10. (8) C B is 151. (9) D C 49.
(10) A D 120. (11) G A is -100. (12) So I got E F and G A to play with to start with. (13) Look at the high/lows. (14) B shows up twice as low. (15) A and G show up twice as low. Second lowest. Yeah. OK.
(16) F... G... OK. OK. OK. OK. So... (17) Forgot to pick it up. Move the G. (18) That didn't do it.
(19) So move the A. (20) That didn't help. (21) Let's move the F. (22) No cigar. OK. No. (beep) OK. (23) Let's go back and try best and start again.
(24) G is on top. (25) Let's go down and move the B around and see where it'll get to with that. (26) Higher. Let's move B again. (27) Oh, OK. So B's gotten in here. (28) Let's move this guy...
(29) Let's move it again. (30) Nope. Wow. OK.
(31) (beep) I understand. I don't like you, but I understand. (32) Hmm. OK. OK. Alright. I accept that. Alright. Alright. (33) What do we do with the
tardy charge? OK. OK. (34) This isn't going to help much. (35) Yeah, I know.
Appendix 6.

Protocol Coding Materials
CODING CATEGORIES SUMMARY

Strategy Formulation:

1) Method Statement
2) Specific Move Description
3) Calculation
4) Observation
5) Information Selection
6) Prediction/Knowledge of System Behavior
7) Explanation of System Behavior
8) Investigation/Questioning of System Behavior
9) Domain Knowledge Statement/Question
10) Trial and Error Statement

Strategy Execution:

11) Performance Monitoring
12) Progress Monitoring
13) Interruption
14) Execution Procedure
15) Operating Help
16) Strategy Help
17) Forgetting

[ 18) Miscellaneous --- All uncoded statements ]
CODING METHODOLOGY

1) A statement does not necessarily fall into any of the coding categories. A statement which is not assigned to a coding category is considered miscellaneous.

2) A single numbered statement can be assigned to more than one category, but the same phrase (words within the statement) should not be assigned multiple codings. Indicate the phrase within the statement which is being coded.

3) Multiple consecutive (or nearly consecutive) statements pertaining to the same episode should be coded once. For example, one specific move (Category 2) can be described in multiple numbered statements. Indicate the statement boundaries of coding episodes extending across multiple statements.

4) Consecutive phrases/statements within the same category but conveying different information are coded as separate episodes. For example, "A is the highest and G is the lowest" is a single statement containing two instances of observation (Category 4).

5) One episode can be embedded within another. For example, phrases coded as an interruption (Category 13) often extend across several lines. A number of separate problems may manifest themselves within the same episode.

Example: 1) What did I just do?
2) I should have...
3) (beep) Oh, let's see.
4) Wrong button. Push space bar.
5) I should have moved job A.
6) OK, so move A.

Lines 1 through 6 are concerned with the main problem of correcting a previous move. Lines 3 and 4 are concerned with the subordinate problem of responding to a keystroke error. Hence, these lines contain two instances of interruption (Category 13).
6) When a subject uses the past tense but seems to be indicating what they are doing now, code it as if it were in the present tense. If the subject is merely commenting on their behavior in a previous session, the statement is considered miscellaneous.

CODING CATEGORY DEFINITIONS

1) METHOD STATEMENT

A method statement describes the way decisions are being made in terms of what is being done and/or why. It suggests that actions are organized and purposeful.

This category does not include describing specific moves (Category 2) or specifying what screen or menu item is being used (miscellaneous).

a) sequencing rule: suggests that actions are part of an organized series. Examples: "I'll do ___ before ___"; "I'll do ___ later."

-does not include "do ___ now"; "do ___ first"; "do ___ next" because something always has to be done now/first/next. The actions described, however, may qualify as a summary statement (see below).
1) **METHOD STATEMENT** (continued)

b) **repetition/stopping rule:** suggests the existence of criteria for determining when to repeat/stop a series of actions. Examples: "do ___ until ___"; "I’m not finished because ___"; "repeat ___ because ___." Includes "do ___ again" when the subject is referring to a process to be repeated.

-**does not include** "do ___ again" when the subject is simply moving a job again or referring to a screen again.

-**does not include** "stop ___ because it’s not working"; "stop ___ because that’s the best I can do" (see Category 11).

c) **specific information search:** a specific item is mentioned **prior** to stating the value or result of a search. Example: "I want to see ___" where the item specified is **more** specific than an entire screen. (When the result or value is stated immediately, see Categories 4 and 5).

-**does not include** total cost information (see Category 11).

-**does not include** simply mentioning an information item (as when reading a column heading).

d) **rationale/goal:** indicates why an action is/is not appropriate, possibly mentioning a goal. Examples: "I’m trying to ___"; "I’m going to ___ because ___."

-**does not include** reasons/goals associated with total cost, including restoring best/previous schedule (see Category 11).

-**does not include** reasons for individual moves (see Category 2).
1) **METHOD STATEMENT** (continued)

e) **summary statement:** paraphrases what is being done/will be done in terms of the information to be considered (more specific than an entire screen), the types/sequence of actions to be taken and/or the subset of moves to be considered. The method need not be a good one.

- **does not include** summarizing what has already been done (see Category 12).

2) **SPECIFIC MOVE DESCRIPTION**

A specific move description indicates that the move being made/contemplated is not completely arbitrary. A move is considered specific when attributes of a job involved, the magnitude of the move, and/or the direction of the move are mentioned (sometimes in nearby statements). A move is also considered specific when a goal/reason is provided. When a particular move description is repeated in close proximity to its first occurrence, it is coded as rehearsal (see Category 14).

In the absence of any of the information described above (attributes, magnitude, direction, rationale), this category **does not include** statements which merely indicate that one job is to be located above/below another, because any move entails putting a job above/below another. Also **not included** are statements which merely indicate that a job is to be moved back to its original position following a cost increase.

a) **attributes of the jobs involved are mentioned:** the attributes are often mentioned as part of an observation (see Category 4) or information selection (see Category 5) episode in a prior statement. The information selection must have mentioned only a subset of job names (reading the order of a schedule does not indicate that the subject’s attention was focused on any particular subset of jobs prior to a move).
2) **SPECIFIC MOVE DESCRIPTION (continued)**

Examples of subcategory a:

"A is above D. I'll move A."
"Job A is really late. I'll move A above G."
"C,D,B. I'll switch C and F."

b) **magnitude of the move is mentioned**: does not include switching two jobs since the jobs may not be adjacent. Examples:

"I'll move it two positions."
"I'll skip one."

c) **direction of the move is mentioned**: usually the words "up" or "down" are mentioned, or in some other way it is possible to determine the direction in which a job is being moved. Includes "I'll put A first/last."

---

3) **CALCULATION**

A calculation statement indicates that the subject has combined numbers to derive a new quantity. Simply reading a series of numbers is insufficient to qualify as calculation (see Category 5).

---

4) **OBSERVATION**

An observation statement presents a conclusion which required the integration of at least two information items, the observation of a pattern over time, or counting of information items. Includes statements concerning the value of information and statements concerning the feasibility of a move. When a particular observation is repeated in close proximity to its first occurrence, it is coded as rehearsal (see Category 14).
4) OBSERVATION (continued)

Examples:

"A has the highest tardy charge."
(unless read from the High/Low Values Screen)
"A is very tardy."
"SPT is the lowest cost schedule."
"This won't help me."
"I can move two jobs around the 256."

-does not include reading items from a screen (see Category 5).

-does not include conclusions regarding total cost (see Category 11).

-does not include stating facts/knowledge independent of the computer system (see Category 9).

-does not include summarizing what has/has not already been done (see Category 12).

-does not include simply mentioning the words "highest", "best", etc. without indicating that the adjective has been associated with a particular value or job.

-does not include reading the jobs/values associated with "highest", "lowest", etc. from the High/Low Values Screen (see Category 5).

5) INFORMATION SELECTION

An information selection statement presents information which is available with minimal processing from the information items on the screen or user notes. The statement indicates that the subject is attending to those information items. When a particular information selection is repeated in close proximity to its first occurrence, it is coded as rehearsal (see Category 14).
5) INFORMATION SELECTION (continued)

Examples:

"A has the highest tardy charge."
(read from the High/Low Values Screen)
"A, C, B, E, F, G, D."
"A, G."
"Best net effect is C and D."
"Best net effect is -100."
"12, 17, 31."
"There are some negatives here."
"All positive."
"A is above/below D."

-does not include reading column headings or menu items from the screen (miscellaneous).

-does not include conclusions regarding total cost (see Category 11).

-does not include stating facts/knowledge independent of the computer system (see Category 9).

-does not include summarizing what has/has not already been done (see Category 12).

-does not include simply mentioning the words "highest", "best", etc. without indicating that the adjective has been associated with a particular value or job.

6) PREDICTION/KNOWLEDGE OF SYSTEM BEHAVIOR

A prediction/knowledge statement indicates a subject's attempts (whether successful or not) to predict the behavior of the numbers in the system, including totals, subtotals and net effects (value, location/existence of negatives/positives, location of best net effect, sign, change of value or sign). The statements can be retrospective, such as "I knew that would be the best net effect." Prediction of total cost behavior should imply a degree of certainty beyond "this might help."
7) **EXPLANATION OF SYSTEM BEHAVIOR**

An explanation of system behavior indicates a subject’s attempts (whether successful or not) to explain why numbers (particularly net effect or total cost) behave as they do. This category also includes statements concerning how the system behaves in terms of the relationship between the items on various screens.

8) **INVESTIGATION/QUESTIONING OF SYSTEM BEHAVIOR**

An investigation/questioning statement indicates that the subject is concerned about the accuracy of a number and/or is attempting to determine why a number behaved as it did. Mere expression of surprise does not qualify as questioning.

9) **DOMAIN KNOWLEDGE STATEMENT/QUESTION**

A domain knowledge statement/question indicates a subject’s attempts to apply domain knowledge to the problem, where domain knowledge is defined as the content of the job scheduling lecture presented prior to the session. Because all of the information items available from the system were explained during the lecture, this category includes statements such as "I don’t know what that means."

This category does not include questions concerning the details of system operation (see Categories 13 and 14).
10) REFERENCE TO TRIAL AND ERROR

References to trial and error explicitly indicate that a subject is guessing, does not know what information is required, and/or has no structured plan for action. This category includes statements such as "I don't know what to do" or "I don't have a reason."

This category does not include statements which do not explicitly indicate trial and error as described above. The following types of statements would not qualify as trial and error: "let me just try it"; "let me go play with it"; "I'm going to mess with it." This category also does not include expression of lesser degrees of uncertainty, such as "I'm not sure what to do" (see Category 13).

11) PERFORMANCE MONITORING

A performance monitoring statement concerns total cost performance (current, previous or best). Examples:

a) stating the value:

"Best achieved is 6800."
"6318."
"Back to 7000."

b) increase or decrease: Usually includes restoring previous or best, unless the statement indicates that the action is an example of a system-interaction procedure (Category 14).

"That worked."
"No good."
"This is higher."
"This is my best."
"I'll restore my previous/best."
11) PERFORMANCE MONITORING (continued)

c) comparison to standard:

"I think this is the best."
"It was lower on Saturday."
"I'd like to get below 7000."
"I'm sure I can do better."

d) overall performance:

"This just isn't working."
"Costs keep going up."
"Nothing seems to help."
"Let's see what will happen."
"Let's see if this will work."

-does not include "Let's see," which may not be concerned with total cost.

-does not include monosyllabic "nope" or "yes" unless accompanied by some of the information specified above.

-does not include investigating/questioning why an increase/decrease has occurred (see Category 8) or explaining why it occurred (see Category 7).

12) PROGRESS MONITORING

Progress monitoring statements indicate what has/has not been done so far or indicate similarities/differences between the current and previous schedules. Examples: "I just finished ___"; "I haven't moved ___"; "I've tried ___"; "I'm back to A C E D."

This category does not include narrating actions being taken (miscellaneous) or statements such as "next I'm going to ___" (miscellaneous or Category 1).
13) **INTERRUPTION**

Interruptions indicate that strategy execution is being temporarily slowed or halted. The delay may be prompted by the belief that a previous statement or action may need correction, by a subject's difficulty with progress monitoring, and/or by the subject's uncertainty as to **how to proceed**. Most episodes of self-questioning fall into this category, unless the question is concerned with domain knowledge (see Category 9) or system behavior (see Category 8). This type of episode often extends across several lines as the subject identifies a problem and determines how to respond.

a) **belief that a previous statement or action may need correction**: includes any (beep) which is acknowledged in some way by the subject. The subject may also discover/indicate the possibility of an error that does not involve operator use (the beep). Examples:

   (beep) "Oops."
   "No..."
   "Wait a minute..."
   "Hold on..."
   "I should have..."

b) **difficulty with progress monitoring**: includes forgetting concerned with actions taken or decisions made during the current session.

   "What am I doing?"
   "Where was I?"
   "What's going on?"
   "Have I tried that already?"
   "What was that?"
   "I forgot to restore previous."
13) **INTERRUPTION (continued)**

c) **uncertainty as to how to proceed:**

"...right?"
"What was that I was supposed to do?"
"How do I ___?"
"I think I'm supposed to ___."  
"I'm pretty sure that I should ___."  
"I believe that I should ___."

- **does not include** uncertainty about cost outcomes, such as "I think this will help" (see Category 11)

- **does not include** uncertainty about domain knowledge (see Category 9), only about what to do/how to do it.

---

14) **EXECUTION PROCEDURES**

Execution procedures are concerned with implementing the decision making process so that processing and memory load, effort and error rates are controlled.

a) **external memory:** creating or referring to written notes.

b) **rehearsal:** repeating information a second, third... time in close proximity to the first.

- **does not include** repeating a method statement, which often appears to be a means of clarification.

c) **double-checking:** repeat actions to avoid/correct errors.

d) **effort allocation:** consideration of costs (time, effort, etc.) and benefits (accuracy, performance, etc.) of a course of action.
14) EXECUTION PROCEDURES (continued)

e) **system-interaction**: actions which are intimately related to characteristics of the system such as its speed or operator behavior. Sometimes actions are taken because the user is uncertain how the system operators behave and wants "insurance."

15) OPERATING HELP

The subject actually requests operating help (not merely discusses it).

16) STRATEGY HELP

The subject actually requests strategy help (not merely discusses it).

17) FORGETTING

Forgetting statements clearly indicate that a subject is attempting to remember, having difficulty remembering, has forgotten, or may have forgotten desired procedural information **from a previous session**. Usually, the words "remember" or "forget" are used.

This category does not include forgetting concerned with actions taken or decisions made during the current session (see Category 13).
### EXAMPLE OF CODED PROTOCOL -- Transcript 5

<table>
<thead>
<tr>
<th>Lines</th>
<th>Category</th>
<th>Phrase Reference/Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>1</td>
<td>keep going until I don't find any negatives and do the net effects (repetition/stopping)</td>
</tr>
<tr>
<td>2-2</td>
<td>13</td>
<td>Oops. I went too far. (unspecified problem, probably use of arrow)</td>
</tr>
<tr>
<td>3-3</td>
<td>5</td>
<td>don't have any negatives left</td>
</tr>
<tr>
<td>3-3</td>
<td>1</td>
<td>play with the holding and the tardy (summary of information to be used)</td>
</tr>
<tr>
<td>3-3</td>
<td>1</td>
<td>to move them more than one (summary of moves to be considered)</td>
</tr>
<tr>
<td>4-4</td>
<td>4</td>
<td>E is tardy alot</td>
</tr>
<tr>
<td>5-5</td>
<td>2</td>
<td>try to move that [E] back (mentioned E in previous line)</td>
</tr>
<tr>
<td>6-6</td>
<td>11</td>
<td>I think this is the best I came up with last time (comparison to standard)</td>
</tr>
<tr>
<td>7-7</td>
<td>1</td>
<td>try things and if [] change it back</td>
</tr>
<tr>
<td>7-7</td>
<td>11</td>
<td>[if] they don't work</td>
</tr>
<tr>
<td>8-8</td>
<td>5</td>
<td>B it's got a tardy charge...</td>
</tr>
<tr>
<td>8-8</td>
<td>4</td>
<td>...but the holding charge isn't very big yet.</td>
</tr>
<tr>
<td>9-9</td>
<td>2</td>
<td>move that up (direction)</td>
</tr>
<tr>
<td>10-10</td>
<td>11</td>
<td>Or maybe not. (experimenter recalls that this comment was made in response to a cost increase)</td>
</tr>
<tr>
<td>11-11</td>
<td>14</td>
<td>write down the order they're in (external memory)</td>
</tr>
</tbody>
</table>
## EXAMPLE OF CODED PROTOCOL -- Transcript 5

<table>
<thead>
<tr>
<th>Lines</th>
<th>Category</th>
<th>Phrase Reference/Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>11-11</td>
<td>1</td>
<td>make sure that I don’t have any really high changeover (summary of actions/information use)</td>
</tr>
<tr>
<td>12-12</td>
<td>4</td>
<td>This one changeover is high from D to G...</td>
</tr>
<tr>
<td>12-12</td>
<td>4</td>
<td>I don’t think it’s any better anywhere else</td>
</tr>
<tr>
<td>13-13</td>
<td>11</td>
<td>That’s better (assumed to refer to the total cost following a change)</td>
</tr>
<tr>
<td>14-14</td>
<td>12</td>
<td>I made some more changes.</td>
</tr>
<tr>
<td>15-15</td>
<td>1</td>
<td>see if I have any negative net effects (specific info. search)</td>
</tr>
<tr>
<td>16-16</td>
<td>18</td>
<td>Narrating</td>
</tr>
<tr>
<td>17-17</td>
<td>14</td>
<td>if I work with them too long (effort allocation)</td>
</tr>
<tr>
<td>17-17</td>
<td>11</td>
<td>...and I don’t make any progress</td>
</tr>
</tbody>
</table>
Appendix 7.

Modified Protocol Transcripts
### KEY TO MODIFIED TRANSCRIPTS

<table>
<thead>
<tr>
<th>Transcript Number</th>
<th>Treatment Group</th>
<th>GEFT Score</th>
<th>Delay until Protocol Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Strategy</td>
<td>11</td>
<td>5 days</td>
</tr>
<tr>
<td>3</td>
<td>Strategy</td>
<td>14</td>
<td>2 days</td>
</tr>
<tr>
<td>4</td>
<td>Strategy</td>
<td>6</td>
<td>4 days</td>
</tr>
<tr>
<td>6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Strategy</td>
<td>9</td>
<td>2 days</td>
</tr>
<tr>
<td>7</td>
<td>Strategy</td>
<td>15</td>
<td>9 days</td>
</tr>
<tr>
<td>11</td>
<td>Strategy</td>
<td>1</td>
<td>3 days</td>
</tr>
</tbody>
</table>

<sup>a</sup>This subject was absent the day of the lecture on job scheduling.
MODIFIED TRANSCRIPT 1

(1) Yeah, I'm trying to remember. OK. (2) I think I go to the "M", Modify Schedule. Right. (3) Analyze the schedule. (4) And the "E" for evaluate net effect. (5) OK, and you push the "B" for the best. (6) Now you write it down, and you go back to the --what the heck was it?-- modification screen, the "M". (7) Yeah, OK so now come down and move the "A" over the "G." (8) And now you go back to the analyze schedule, and you start the process all over. (9) With the "E", up, "E", up, "E", up. (10) We know what that is. (11) Oh, wait a minute--no, we don't. OK. (12) What I noticed on all of these was that we had one letter on the top that kept moving down. (13) Back to "A." (beep) Oh, OK. (14) Back to modify. Sorry about that. (15) Now this is where it got confused, because it went up. (16) And I believe I go back to the "A" and I continue to do this same thing. (17) You just continue to repeat this until you get it right. (18) I don't know. (19) Push finished. (20) Although it seemed-- I thought I had a lower score than that Saturday.
MODIFIED TRANSCRIPT 3

(Accidentally recorded over first few moments of protocol)

(1)...and I'm going to go back. (2) I see that I've made an improvement, and I'm going to go back to the analyze screen and continue this until I get rid of all my negatives. (3) I really don't even need to do these again because I knew I knew that was a -34, but I have to do all of them because the top 2 are going to change. (4) Might be... looks like -74. (5) OK. So I'm going to press "B" for the best which was -74 and switch those 2 jobs. (6) Going to the modification screen. (7) D is to go above G. (8) OK. Go back to analyze. (9) Still have a negative, so I know I have another round of this. (10) Just moved another job. (11) Going back have to analyze again the schedule. (12) I still have negatives. (13) Going to modify B above G. (14) Go back again. (15) I'm pretty sure I still have negatives. Yep. (16) But I can see my numbers becoming larger into the positives. (17) OK. This should be my last one. (18) F before E. (19) But I'm going to analyze the schedule one more time to make sure. (20) Positive, positive again, positive, positive, positive, positive. (21) OK. So
I’ve now rearranged the jobs and I have I have all positive net effects and the next step is to look at changeover costs. (22) I just got done concentrating on inventory and tardiness. (23) And I’m looking for the highest changeover cost which is 300 and I’m going to change jobs that are around the 300. (24) I’m just going to jot this down here. (25) I’ve got 300 is the largest, D, A, C and B and I can basically make... (26) Let’s see, I can move this... I can make 3 moves here around that largest changeover cost. (27) I’ll do one at a time and go back to... Oops. Hang on a second. (28) I need to I need to do something here. (29) I need to go back to the modification screen before I proceed and I need to restore my best schedule before I continue on. (30) OK. Now I can begin moving these jobs. (31) I’m going to start by moving in small increments so not to make such drastic changes in the tardiness and inventory cost as I change the changeover costs. (32) OK. My previous was better but I’m still going back to my previous because I have to calculate the other ones. (33) I’m going to move C next. (34) I did better this time, but I still have to go back to previous and do one more change in the job arrangement. (35) I’m going to take D above A. (36) OK. It looks like that one came out the best, so I’m going to
restore that. (37) I don’t know if the system
discovery automatically does that so I’m just being safe here.
OK. (38) My numbers now my changeover charges in that
area have gone down into the 100’s. (39) I do have
one large one, 249, so I’m going to attack that next.
(40) And I can only make 2 changes here I believe
because it’s towards the bottom. (41) Start by moving
this job here. (42) I didn’t do better. (43) I got
worse but I’m still going back to previous and then
I’ll move the next one up. F above G. (44) And that
got worse, so I’m going to get my previous up. (45)
I’m going to just hit restore just in case. (46) I
just want to make sure it restores that as the best
schedule. (47) So there’s really I can’t get rid of
the 249. (48) I can go to the next highest changeover
cost which is 197 and I can make some changes here.
(49) I can start by moving F above G. (50) I did not
do better. (51) Go back. I can move... (52) Oops.
See now I did that move twice now that I look at my
sheet. (53) I already did that. (54) G. I did not
move G, though. (55) Let me try moving G. (56) I
did not do better. (57) I’m going to write these down
now so that I don’t do that again where I did the same
calculation twice. (58) And B to C I already did so I
don’t need to do that again and the next one is 136.  
(59) I’ve already made E.  (60) I’ve already changed  
F.  (61) Let’s see, 114.  (62) I’ve already changed  
most of them I think.  E,F,G,B,C,D,A.  (63) So I’ve  
really completed the short moves with the changeover  
costs so next thing I can do is make double moves.  
(64) So I’m going to go back down to the 249 which is  
the largest changeover cost.  (65) OK. So I’m moving  
to E.  (66) There’s two moves there.  (67) And I  
didn’t do any better.  (68) Try jumping the next one  
up two moves.  (69) And I did not do better.  (70)  
I’m looking at the totals which is really where I’ve  
paid most of the attention the whole time.  (71)  
Moving the next one up two.  (72) Oops. I think I  
made an error here with the arrow.  (73) I’ll see when  
it comes back.  (74) Yeah. I accidentally didn’t skip  
two jobs.  (75) I went one.  (76) So let me go back  
and do that again.  (77) I should have gone above the  
A.  (78) Nope. UhUh. Not making any improvements  
here to the schedule.  (79) What I should have done,  
though, is-- I went all the way up-- I should have gone  
back and still stuck near the highest numbers and came  
down.  (80) There’s up movements and there’s also down  
movements with the jobs.  (81) So I’ll go back down to  
where I should’ve...  (82) (beep) Oops. Wait. Do I hit
escape?. OK. (beep) Hey, why can't I do that? Oh, OK. The wrong arrow. The wrong way. (83) I'm back doing the double jumps but I'm going down versus up and I'm not seeing any improvement. (84) I've got another one here to do. (85) Nope. Basically I think this is it. (86) I've made single movements. (87) I've made double. (88) I'm not going to do triple because basically I think the further you jump the less improvement that you see so I think that's why I stopped there. (89) Should I continue to another job?

EXPERIMENTER: If you're done with this job set, yes.

OK.

(90) I think I've seen this number before so I'm fairly certain that I've completed this. I'll just say 6. I'm not totally. (91) And my strategy I'll give a 6. There's room for human error.
MODIFIED TRANSCRIPT 4

(1) OK. Now I don’t know if I cheated when I did this but I did not use this schedule here. (2) I went on to the next page. (3) Well, modify the schedule. (4) And I went directly to analyze schedule. (5) OK. Now I use the net effect and the I&T. (6) "E." (7) Do I have to talk this out even as I go here?

EXPERIMENTER: If you're thinking say what you're thinking.

(8) Oh, OK. I'm not thinking really. (9) Now I press "B" for best net effect and I move on to... (10) I want to see the best net effect for B. (11) Oh, OK. I did that. (12) Compare Schedules...Now I go to...I forgot how to do this. (13) I modify screen with the best. (14) OK. The highest number is 300 and I use the first job above it and I move the job out and I go up one and place it here. (15) And this is the best one the best achieved. (16) And, let's see, it changes my schedule so now I go back and I analyze schedule. (17) And I do that I&T. What is it E? Evaluate. (beep) Oh. E. (18) I forgot this is slow so I have to take my time with it. (19) Go with the best and now I'm going to modify... (20) This is the best I want modify. (21) The largest number is 271 so
I'm going to go down to B which is E which is one above take it out and move it up one. (22) OK. Go back to the best one. (23) The largest. OK. (24) Go one below. (25) Take it out. (26) Move it up one.
(27) This is the best achieved. (28) Analyze schedule again with the new best. (29) "E." I'm hoping I'm doing this right. Best. (30) Now I have to do it again. (31) Modify. (32) Largest number is...I'm repeating myself here. (33) I'm writing down what the largest changeover charge is. (34) And I'm going to start moving jobs. (35) Taking the job above it and I'm moving it up one. (36) That's too high so I'm going back to the best achieved. (37) I'm going to the largest number the next largest number is 216.
(38) G is above. (39) I can't move that. (40) D is under it. (41) I'm not sure if this will change it but I'm moving it above. (42) It changed it to a lower grand total so I'm going with the best for this one. (43) Now I'm going to the next highest number which is 258. (44) The job above it I'm going to take out and move it up one. (45) That was too high so I'm going back to the best. OK. (46) I'm going to the job below, taking it out and moving it up one.
(47) That's too high, so I'm going back to the best.
(48) OK the next highest number is 144 which I'm writing the numbers down. (49) I tried D and G with 256 so I'm going to take the letter the job below it and I'm going to move it up one. (50) That's lower, so I'm going to use this as my best achieved. (51) 140...Umm. I'm losing track of what I'm supposed to be doing here. (52) I'm starting over with my moves because I think I messed up somewhere. (53) So I'm going to the highest number changeover and move my jobs. (54) The one above. (55) Going back to my best. (56) One below the highest job one above. (57) Go back to my best. (58) The next highest job is 162. (59) Writing this down so I won't forget. (60) The one above and move it one position up. (61) Go back to my best. (62) One position down. (63) Go up one. (64) Going back to my best. (65) Two positions down up one. (66) Going back to my best. (67) I'm going to my next highest number which is 158. (68) One position above goes up one. (69) Going back to my best. (70) 158 up...moving it up one position. (71) Going back to my best. (72) OK. Two jobs down. (73) Taking it out and moving it up one position. (74) Going back to my best achieved. (75) My next highest number after 158 is 140. (76) D can't go up any higher. (77) A is below it so I'm going to take A out
and move it above D. (78) I don’t know if this is
going to change it but I’m going to try it. (79)
That’s higher so I’m going back to my best. (80) Two
positions down is G. I’m going to take that out and
move it up one. (81) That’s higher so I’m going back
to my best. (82) And the next highest number is 126.
(83) I did A with 140. (84) I did G with 140. (85)
And I did C with 162 so my best is 6029 so I’m going to
hit “F” for finished. (86) Are you completely
finished making changes to the schedule? Yes. (87)
Using the 7 point scale shown, how sure are you that
you found the lowest cost schedule possible? I’m not
very sure that I found it. I’m going to say about 4.
(88) Using the 7 point scale— I know I forgot a st—
I don’t know. Can I go back? I forgot to move
positions twice for the lowest. I just remembered.
How sure are you that the strategy...I know I skipped a
step so I’m going to hit 2 here. Are the values...yes.
MODIFIED TRANSCRIPT 6

EXPERIMENTER: Try to talk out loud.

(1) OK. Well, I remember from Saturday that I have to
go through to find the lowest cost so I’m going to...
(2) This has already been analyzed so... (3) I’ve got
to play with it for a minute because I don’t exact...
(4) OK I’m going to analyze the costs so...
(5) I’m going to evaluate the net effects. (6)
(bEEP) Oops. Oh. Now I want to see which grouping
gives me the best net effect. (7) OK that was the -
100. (8) Let’s see. Got to return to the... Oh.
Going to return to the modification screen. (9)
Change G and A. (10) I modified the schedule by
switching G and A around which gave me a better cost
than what I previously had. (11) OK. Let’s see.
(bEEP) Oops. OK. Let’s see. Do this one. (12) Oh,
now I think I’ve forgotten to go back to my previous
screens. (13) That blows my other... I want the best.
(14) OK. Now. OK. Ooh. That’s good. (15) Well the
best I’ve achieved is 6029. (16) I think that’s the
best. (17) Oh. That wasn’t what I wanted to do. (18)
(bEEP) OK. Now how do I get out of here? Yes.
MODIFIED TRANSCRIPT 7

(1) OK. Analyze the schedule comparison. (2) OK. Smith’s Rule is out of order. (3) Shortest Processing Time, Weighted Slackness Rule, Least Changeover Heuristic and Random Order. (4) Looks as though shortest processing time and weighted slackness rule are very close in total cost and random order is the highest, hmm. (5) Modify. OK. (6) Oh, least changeover heuristic is the lowest. No, it’s not. (7) Let’s see. Shortest Processing Time. OK. Modify schedule. (8) OK. I get this chart the shortest processing time modification screen. (9) What I’m going to do here is go down. (10) No. Wait a minute. I’m not going to do anything. (11) I’m just going to look at this and then I’m going to go to another screen, right? (12) Wait a minute. Analyze the schedule. That’s what I want to do. (13) Restore, job to be moved, new location, delete, compute new costs, restore previous, best achieved. Let’s see. What was I doing? (14) Oh, I was making changes. Help. Help. Operating Help. (15) About the information...to view additional information before making a change press "A" for analyze. Press page down for the next help screen. (16) (beep) Hmm. That’s it. End. (experimenter turned off the numlock which
was on) Oh, OK. Thanks. (17) Use the up and down
arrows to highlight the job... (18) Leave help. (19)
OK. "A." I want to analyze the schedule. (20) OK.
Best net effect is highlighted. Hmm. (21) Evaluate.
Hmm. Evaluate. Evaluate. Evaluate. Evaluate. (22)
Now, best net effect was that one. OK. G and A.
(23) Return to modification screen. (24) Let's see
if we can get better. (25) Let me see. G and A. I
won't forget. OK. G and A. (26) What am I doing?
I'm going to make G and A G and D. (27) Let's see.
Take that. Enter. (beep) Oh, no. Left arrow. (28)
That one I want to put up there. Enter. Calculating
costs. (29) Oh, I forgot to look at the cost. Let's
see. (30) Best achieved. Previous total. Oh, that
came out better. 6368. (31) Well. I did G and D
instead of G and A. (32) Now, that's it. Go back to
the other one. (33) Analyze the schedule again.
(34) Evaluate. Evaluate. (35) Hmm. Best net effect
is now D and A. (36) Modify it. (37) I'll make that
A and D or D and A again. (38) Back to 6368. Hmm.
(39) Back to G, A and D. (40) I'm going to analyze
it again. (41) Oh, no not analyze this. Oh, wait a
minute. What did I do? No, that was right. (42) OK.
Analyze the schedule data. (43) OK. What am I doing?
I&T net effect. (44) Best computed net effect.
That’s what I’m looking for. (45) Now...inventory.
Ooh, wow. Seems like more periods are later now so.
(46) Evaluate. Evaluate. Evaluate. (47) Back to
where I started. Hmm. (48) Modify. No wait. What
was that? (49) Analyze again. (50) Now what was the
best one? (beep) Oh, shoot. (51) I have to do it
again. But it’s G and A. (52) It’s G and A. I knew
it was. OK. (53) I’m going to modify G and A again.
(54) Forget all that. I’m going to make... (55) I’m
going to do it this way. (56) I can do it any way I
want to? I’m going to make F go up here. See if I get
anything lower. (57) Hey, 6368. Same thing. (58)
That wasn’t anything. (59) "P" to restore the
previous. Same thing. Hmm. (60) Analyze the
(62) G and A again. OK G and A. (63) Modify. (64) G
goes here. A goes there. Calculating cost. (65)
Hey, it’s lower. OK. (66) Go back. 6317. (67) Now,
evaluate. (68) Best is G and D. (69) Now, modify
it. G and D. (70) We’ll put D here and G right
there. (71) OK. 6187. Great. Now, moving right
along. Analyze. (72) OK. The best. G and C. OK.
(73) Modify G and C. (74) Take C and move it above
G. Enter. (75) Now they’re just about all...Nope,
they're not all negative periods late. (76) Let me see. Duration of the job: 4, 5, 8, 3, 9, 10, 10. (77) This isn’t telling me anything. Let’s see. (78) Evaluate again. (79) Best. (80) Figured. 71. G and B. It’s going right down the line. (81) OK. Modify that. G and B. (82) And take B and move it up above G. (83) Enter and I bet it will be lower yet. Let’s see. Has to be. (84) Yep. 6037. I knew it. (85) Go back to the first screen. (86) Analyze the schedule. (87) And the next one’s are going to be. I bet G and E. Watch. (88) Evaluate. Evaluate. Did I have that? OK. Evaluate. Evaluate. Da dah. Da dah. (89) Best. (90) No. E and F. Hmm. E and F. OK. Modify E and F. (91) We’ll put F above E. Calculating it. (92) It’ll be lower maybe. Maybe. (93) Yeah. 6016. Wow. How about that. OK. Go back to analyze for the last time. (94) Now either G and F or F and E again. No, let’s see. Hmm. (95) Evaluate. Evaluate. None of them. (96) Hmm. Got ‘em all. That’s it. OK. Now (97) We’ve got the best. Hmm. Forget this schedule. (98) Let’s see what it looks like. (99) Everything’s got negative late periods except E and that’s about it. (100) Everything’s negative. OK. Now. Hmm. (101) "M." Modification
screen. (102) I’m ready to do something. What am I doing? I’m going to move two jobs and make them better right? (103) Best achieved is 6016. Great. OK. I’m going to take...Hmmm. (104) Oh, wait a minute. Am I going to move two jobs or am I going to move one job two spaces? I’m going to move... (105) Let me see this thing now. Tardy charge. Hmm. Nothing. Oh, because there wasn’t any tardiness. Periods completed. (106) What is this? The shortest processing time modification screen. (107) OK. What I’m going to do is rearrange these and compute a new cost for the shortest processing time. (108) Oh, oh, I see. That last screen we must have got the shortest processing time. (109) This screen we’re going to fix it even better. OK. Now. Hmm. (110) Best achieved 6016. That’s what I’ve got to beat. (111) So we’ll go...Hmmm. A,D,C,B,G,F,E. (112) What am I doing? C take B. Wait a minute. (113) A, D and C fit together. If we take A...No no no. (114) Take C and move it above A (beep). (115) Oh forget this. Help help help. Let’s see. (116) Next help. Use the up and down arrows, use the left arrow, use the up and down, use the... (117) Next help. I’ve done all that. (118) To undo a change, to indicate that you have finished, press "P," to undo a series of changes, to
indicate that...No, next help. (beep) (119) No next help? Hmm. Previous help. Page up. (120) To make a change, modification screen, along with the total cost, about the information on the screen... OK, wait a minute. (121) Next help screen, along with the total cost information for the schedule currently displayed on the screen, you are also provided with total cost information for the previous schedule, this will only exist if you have made at least one change, any number of, best achieved, this number represents the lowest total cost schedule you have achieved in the course of modifying the schedule displayed on your screen, OK, press "A" for analyze. Oh, (beep) no previous? End. (122) Alright. What am I doing? I'm going to take two at a time. (123) No, I'm going to skip two. (124) I'm going to take B and move it above... D? (125) Let's see. No no. I want to start with A. (126) Delete delete delete delete. Delete you too. Where are you? Uh oh. (127) OK. Start with A. I'm going to take A and move it down below C. (128) Alright. That's taking two at a time. (129) Oh, god. Made it worse. OK, nope. Don't want to do that. (130) Put A back. (131) No, restore. How do you do that? Let's see. Restore the best previous. (132) I'm
going to take D and move it down below B. There.

(133) Nope. Doesn’t work. Restore. (134) Let’s see. Now I’ll take C and move it down below G. (135) Shortest processing time. Nope. Inventory holding charge, tardy charge, period completed. (136) Why would I want to do that? What’s this got to do with it? (137) I did A, I moved D below B, I moved C below G. (138) Let me finish. B below F. (139) Nah, that doesn’t work. Nothing works. F below...Let’s see I did A below C, D below B, C below G and I did B below F and I’m going to take G and go below E. Da dah. (140) Let’s see. Two at a time. (141) Oh, god. Restore. It’s getting worse by the minute. This is not improving at all. (142) OK. 6016. That was the best. OK. (143) I moved them one step and I moved them two steps and they don’t get any better so I am finished. That’s it. (144) You will not be able to return. Nope. Are you completely finished making changes to this schedule. Yes. (145) Wow. Oh, this thing. OK. Absolutely sure. Using the seven point scale shown above, how sure are you that you found the lowest cost schedule possible for the job set you just completed. Please respond with respect to schedules which can be created after attempting 100 or fewer changes. How sure...7. (146) How sure are you that
the strategy you used was the best way to solve the problem... Six. There had to be a better way. That was about the best. That was about the best you could do. Are the values you entered above correct? Yep.
MODIFIED TRANSCRIPT 11

(1) Press "M." (2) "A" for Analyze. (3) (beep) And it's beeping already. Come on think. (4) (beep) It's shift. Unshift. There it is:

EXPERIMENTER: It was locked?

(5) Yeah. Let's see. You want to look for all the negatives. (6) OK. The highest net effect is G A so I want to go to the modify screen so I can switch G to A. (7) The computer is slow. (8) And I want to go back to analyze to see the results. (9) (beep) "E" for evaluate. (10) Go to the best net and it's going to be G G to D so I'll switch those. (11) I'll go back to modify. (12) Analyze. (13) Evaluate. Evaluate. Evaluate. (14) "B" for...I'll switch C and G. (15) Go back to modify. (16) Analyze. (17) "B" for best. (18) I'll switch G and B. Back to modify. (19) Analyze. (20) "E." (21) "B" for best. Switch E and F. (22) (beep). Oh, dingaling, you don't want that. OK, you want to you've already press any key to continue and I want to cancel that so...move, compute new costs, restore, hmmm, move to a new location...I should have a delete key. (23) It's awfully slow. "A" to analyze. (24) "E." Eek. E and F. E and G. (25) Modify. Move E and G. Enter.
(26) Hmm. My best achieved was 60...Well, I’ll ignore that. (27) Go back to analyze. (28) Evaluate. (29) This should be my last negative. (30) Move E and F. Modify. (31) Hmm. That’s even better. (32) OK, let’s go back to analyze and check that I’ve moved all the negatives. (33) Then I’ll go down and just double check that I didn’t miss anything.' (34) Then I’ll go back to modify. (35) OK, Now I can move to the next step of the changeover charges. (36) I’ll start with the largest which is from D to C which is 300 and I’ll switch those. (37) Hmm. That’s better. 5992. (38) Let’s see if I can get rid of the changeover from D to B which is 271. (39) That’s not good. I don’t like that one. I’ll go back to previous. (40) Let’s see if I switch... Let’s try the E switching F and E since that’s my lar-- next largest changeover. (41) No, I’ll go back to previous. (42) Let’s go let’s switch F and G. (beep) Oops, wrong button. (43) Nope, don’t like that. Let’s go back to previous. (44) Let’s see if I do F two places. (45) I don’t like that one. Go back to previous. (46) See if I go to the next the third largest 197 which is the changeover between B and G what happens there. (47) Don’t like that one either. Let’s go back to previous. (48) Let’s go
back to the change between D and B and see what happens
if I move B to C. (49) No, I don’t like that. Let’s
see if I move F...G. (50) I don’t like that total
cost. Let’s move A. (51) No, I don’t like that one.
(52) I feel I have the best cost I have the best
achieved so can I move out? (53) Have you completely
finished making changes to the schedule...yes. (54)
Using the scale above...the lowest cost possible for
the job set you just completed. Please respond with
respect to schedules... I’ll say seven.
Appendix 8.

Generalization Task Protocols
KEY TO TRANSCRIPTS

<table>
<thead>
<tr>
<th>Transcript Number</th>
<th>Treatment Group</th>
<th>GEFT Score</th>
<th>Delay until Protocol Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Strategy</td>
<td>11</td>
<td>5 days</td>
</tr>
<tr>
<td>2</td>
<td>Operations</td>
<td>17</td>
<td>3 days</td>
</tr>
<tr>
<td>3</td>
<td>Strategy</td>
<td>14</td>
<td>2 days</td>
</tr>
<tr>
<td>4</td>
<td>Strategy</td>
<td>6</td>
<td>4 days</td>
</tr>
<tr>
<td>5</td>
<td>Operations</td>
<td>18</td>
<td>4 days</td>
</tr>
<tr>
<td>6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Strategy</td>
<td>9</td>
<td>2 days</td>
</tr>
<tr>
<td>7</td>
<td>Strategy</td>
<td>15</td>
<td>9 days</td>
</tr>
<tr>
<td>8</td>
<td>Operations</td>
<td>5</td>
<td>5 days</td>
</tr>
<tr>
<td>9</td>
<td>Operations</td>
<td>5</td>
<td>2 days</td>
</tr>
<tr>
<td>10</td>
<td>Operations</td>
<td>2</td>
<td>2 days</td>
</tr>
<tr>
<td>11</td>
<td>Strategy</td>
<td>1</td>
<td>3 days</td>
</tr>
<tr>
<td>12</td>
<td>Operations</td>
<td>11</td>
<td>4 days</td>
</tr>
</tbody>
</table>

<sup>a</sup>This subject was absent the day of the lecture on job scheduling.
TRANSCRIPT 1 -- EXTENSION

EXPERIMENTER: OK, just one more job set, but this time there's an extra consideration. Whatever schedule you come up with, you have to make sure that job A is not late. So you're still trying to get a low cost schedule but, at the same time, you have to make sure that job A is not late.

(1) A has to be on time. (2) You don't care anything about inventory holding cost as long as A is on time.

EXPERIMENTER: Well, you do care about total cost.

OK, yeah.

EXPERIMENTER: There are a number of ways you could set it up so that A is on time, but that has to be true.

OK.

(3) I did that wrong. (4) Go back to "A." (5) Interesting. It's already in the best location. (6) OK. It wasn't in the best location. (7) I forgot what I was doing. (8) Back to "A." (9) There's really not much going on in my head except I'm waiting to see which one of these is going to produce the best results. (10) Back to "A" again. Oooh. (11) So actually we've finished, so now we go back to the...woah. (12) Out of curiosity, I'm going to see what "C" does. (13) Whatever. Back to modify. (14) And I assume I want to get the one that restores me the
best one. (15) A cannot be late. Ah, hah hah, right, OK. (16) And there are tardy charges here. (17) OK. Just got interesting. (18) (beep) Alright, alright. How do you do bottom on this thing... (beep) It figures. (19) Didn’t help. (20) I’ll move A to see where it moves around to, just experimenting. (21) Ah Hah! Well, we’ve got A with no tardy charge so we’re going to go to the modify screen, or the "A" screen, whatever. (22) That’s better than it was. (23) Back to "A." (24) What? I have a little trouble believing that it worked that easily. (25) OK. Back to modify. (26) I guess that’s it.
TRANSCRIPT 2 -- EXTENSION

EXPERIMENTER: The difference with the second job set is that whatever schedule you finally come up with, job A cannot be late. Let’s say for political reasons that job has to be done on time.

OK.

EXPERIMENTER: So whatever else, whatever other changes you make, just be sure that job A in the final solution is on time.

OK. Job A. OK. (1) The lowest cost here would be 7173 which again is shortest processing time. (2) Analyze the schedule. (3) Do evaluation of net effect on each one of them starting at the bottom. (4) So, let’s see. (5) Evaluate. (6) Positive 30. (7) Job A can’t be late. (8) Right now it’s very late. (9) OK. 63. 181. There’s a good change—D to C. (10) B to D no change. (11) And there’s a good change. (12) OK. Job A cannot be late and now it’s 25 periods late. (13) I need to change it so it’s not late. (14) So it needs to be done. (15) Let’s see we’ve got 10...

(16) OK. Modify. (17) Job A can’t be late, huh? OK. (18) Might as well figure out how it’s going to affect the schedule. (19) It needs to go right I believe right there. (20) It’s going to increase.

(21) OK. We’ve completed. (22) It’s 25. (23) It’s what it needs to be late—be there. (24) OK. Now I
go back and analyze based on job A not being late.

(25) OK. Net effects. (26) Of course. OK. (27) That’s a possibility. OK. (28) If I move job A up in the schedule, move it before D, that will save me some money. (29) And it did. OK. (30) So if job A saved me some money do some more evaluation here. (31) Analyze the schedule. (32) See if there’s any changes down here. (33) 63 maybe. (34) It’s low enough to where an effect may be. (35) Oh, there’s a good one. (36) Don’t have to go any further. (37) Modify, so D and C. (38) Got to tell it to go. (39) Oh, wow. Interesting. How about that. Best achieved and it’s not late. (40) OK. Do some more evaluation. (41) 63-- I’ll come back to that one. (42) Can still move that down. (43) That’s a negative effect, I might be able to move it back down. (44) Get a positive effect there. (45) OK. Best net effect was that one, B and C. (46) I forget. D and F right? (47) Yeah. Change D and F. (48) Ooh, It went up alot. (49) Go back to best. (50) I can change A and C. (51) I don’t think it will make me late. (52) Let the computer tell me. (53) That should save some funds. (54) No it doesn’t--yeah, it does and I am... late--bummer. (55) Let’s see. Analyze. (56) I was hoping it wouldn’t make me late. (57) I’m a week late. (58)
What can I do to change my week late? (59) If I have D, 6 weeks, this is 9. (60) If I change those 2 what will happen? (61) The quickest way to find out is to do it. (62) Move that to there. (63) Ignore the calculation. (64) And move that to there. (65) And see what happens. (66) Wew, Big cost. I’m not late though. (67) Check analyze again. (68) Forgot which one was the best beforehand. (69) I know it was...
(70) OK. I forgot which one was best so I have to go back and find it. So I don’t forget. (71) So modify. (72) Go back to best achieved which won’t work, but the only difference was that I moved these two. (73) So it was E,B,C,A,D,F,G gives me the best but I can’t do that so it’s E,B,C,D,A,F,G. (74) So far that was the best without it being late. (75) I want to go back and play with what I had before now. (76) So I do my previous list. OK. (77) Evaluate or analyze. (78) See if there’s any net effect changes I can do. (79) 63,53. I can’t move these. (80) I can move these though. (81) OK. I might change there and there’s a change there. (82) So I can move A to D. (83) OK. So move. I can move it up in the schedule. (84) OK. It reduced it. (85) I can move this because that showed a net effect. (86) Let’s see if
that actually works. (87) I've done this before and it didn't, but things have changed. (88) It's worse, so previous. (89) Back and evaluate. (90) And let's evaluate the bottom three. (91) G and F is a positive effect, C and F is a positive effect, C and D is a big negative effect. (92) Let me change that. (93) Modify D and C. (94) Swap D and C and I should be under 7. (95) Yes. 6906. And that is E, B, A, C, D, F and G is 6906. (96) And let's see if I can make this any better. (97) So analyze. (98) The bottom three again. (99) 63 I don't believe will change much.

(100) OK. I might be able to do some more there. (101) That should be a big positive if it'll ever do it. Yep. (102) I can't change these I don't think. (103) Let's see. I'm 8 weeks early. (104) I can't change C and D because I've only got 8 weeks. (105) I have to do C. (106) Sounds like a good idea but I can't do it. (107) Evaluate A and C. Big change. (108) That one I've already done. (109) It's a negative but it doesn't affect it. (110) So modify...B and F. (111) It wasn't much but let's see what happens. (112) Previous change it went up. (113) Analyze. OK. (114) I'm 13 weeks early and I've got a 9 dollar per period charge. (115) That's not too bad. (116) 31 weeks early on a 41 dollar
charge. (117) And it’s telling me that that’s cheaper? (118) That doesn’t—I can’t believe that.
(119) Let’s see. Somewhere there’s a charge. (120) It’s got to be in setup. (121) B to A, B to A is 160.
(122) A to B is 220. (123) There’s a charge there.
(124) E to A is 265. (125) E to B is 215. (126) Where’s it coming up with that difference? (127) It’s too high. (128) 34 dollars, 10 weeks. (129) I’m going to try it. (130) I don’t trust it. (131) It wasn’t lying. (132) Inventory holding cost is high on C. (133) If I change C with D, I might reduce because of inventory holding cost but charges might be higher.
(134) Yep, it is. Previous. (135) No savings by changing there. (136) If I look at F, 16, and G, I swap—tardiness charges. (137) I’ve done it I believe, but let me just try it. (138) I wish I had an AT class machine. (139) It would be alot faster.
(140) Hey, how about that? That worked. (141) OK. Swapping F and G definitely helped. (142) So E,B,A,C,D,G,F is at 6784. (143) That was an improvement, and I believe that was an improvement over my first one of E,B,C,D,A,F,G. (144) Pretty close to best achieved, and is A late? No. (145) Go back and evaluate. (146) Let’s see if there are any changes
that will help here. (147) That should be a negative, OK a possibility. (148) That’s a po--still a negative! (149) That’s what I just changed, but maybe, maybe. (150) Positive, yes. (151) I can’t change these two although I’d like to. (152) 206. And I’ve already done that one. (153) The best effect was, of course, that one. (154) OK. Let’s see what changing D and G at this point does. (155) I’ve moved them out of order a little bit. 6784. (156) 68. It’s higher. (157) Go back to previous. (158) OK. That’s as good as I can get on that one. (159) Let me go back to, let me make sure I’ve got it written down right: E,B,A,C,D,G,F. OK. (160) I wish there was a faster way to get back. (161) Say move this, this, and this. OK. (162) Best achieved...and A is late. (163) And the only difference with that is that was those two were changed, or was it? (164) Serves me right for not writing it down. (165) OK it was 6906. (166) OK. So that’s not better; so I need to change it all back to the way I had it. (167) E,B,A,C,D. Oh, that’s neat. (168) That takes it down to 784. (169) Analyze one more time. (170) 13 weeks early, 31 weeks early, doesn’t affect, 8 weeks early so A is on time OK, C is late by 2, D is early by 5, inventory cost is 6, tardiness is 50, so B early. 5 and 43, 24
periods late. (171) That’s it. That’s the one.
(172) Modify and I’m finished. (173) And that’s as
good as that one’s going to get and I feel pretty good
about the way I did it.
TRANSCRIPT 3 -- EXTENSION

EXPERIMENTER: OK, now this next time through... Previously it didn’t matter what order the jobs were in. Now this next time through, let’s assume that for some reason the people who’ve given us job A absolutely must have it on time. So in making the schedule you have to make sure that, no matter what else you do, job A is not late.

(1) OK. This screen here doesn’t tell me how many days. (2) It shows me the total, the charges basically so... (3) Let me see if the analyze schedule screen will show me more information on that. (4) Duration of A is 10. (5) Periods late. Now, can I ask you a question?

EXPERIMENTER: Sure.

(6) The negative was...

EXPERIMENTER: Early.

Early, and the OK. (7) So right now we’re 25 days late the way we’re scheduled so we have to do something about that. (8) I should I’m going to just continue with the same approach and see if I can improve the periods late. (9) Ooh. I didn’t even get a negative there. (10) I thought I would. (11) That’s not going to do anything here. (12) I’m basically going back and doing the same thing I have done, making sure
I don’t have any negative effects in here, and then I’ll concentrate on job A after I clean up some of the rest of the schedule. (13) Go back to analyze. (14) Ooh. Just thinking now. (15) If I do this and then I move A, I might really mess the schedule up. (16) Still 25 days late. (17) If I switch it with the job right before it that’s 23 days I’m going to gain. (18) I think I’m going to make that switch now, and then go back and do negative effects. (19) Uh oh. Wait a second. I just made it worse. (20) No, I didn’t. I’m looking at the wrong screen. (21) No, I’ve improved it. (22) I’m still 16 days late, so I still need to move it up. (23) Can I make any larger jumps? On this screen? (Experimenter replies) Just on the other screen? (24) Well, I’m going to move job A one more time. (25) And analyze the schedule. (26) It’s still late. (27) Duration of job 10. (28) I’m going to do one more modification. (29) Bring it up one more job. Before F. (30) I should have moved it up in larger increments. (31) Still late. (32) I’m moving job A one more time. Hopefully, the last. (33) I’ve got job A in a location now where it’s not going to be late but I don’t know what I’ve done to the rest of the schedule so I have to look at that. (34) OK
there. Ooh. That's going to put me right back where I was before. (35) I'm going to try moving A again. (36) Well, by moving A I just created... (37) I guess I have to make a tradeoff here. (38) The duration of A is the longest of all the jobs. (39) I have to complete it on time. (40) Normally I would attack the shortest jobs, those taking the least amount of time first. (41) Not even looking at changeover costs at this time. (42) If I take A and E and swap them I'm going to make it even worse I think. (43) I'm going to try that and see. (44) OK. That worked. (45) I have job A on time. (46) Actually, it's going to be completed 17 days early, and I have no negative net effects, so I'm going to continue on at this point. (47) I'm going to go to the other screen and start looking at the changeover costs. (48) I guess when you have a computer to do these kind of calculations for you, sometimes you don't think it through as thoroughly because you can hit a few buttons and get answers a lot quicker than if you actually sat down and had to do it from scratch. (49) Sometimes I think we lean on our computers, I do anyway, to do a lot of the calculations for us. (50) OK. The 296 is the highest changeover cost and I'm going to make 3 single jumps. (51) This is now I'm back on track to our original instructions
that you gave. (52) D one move. Let’s see if I improve the grand total by reducing it. (53) Yes, I have. But I’m going to go back to previous because I might be able to do better by moving F. (54) That time it got worse. (55) I’ve got one more move on this one...and that would be C. (56) Nope. I’m going to restore the best achieved schedule which is 6553. (57) OK. It looks like I can make some more changes now because I still have the 296 which is a high number. (58) I don’t want to move D back to there because I just flipped those two. (59) Nope. Nope. It’s getting worse. OK. (60) Well, let’s go to the next highest number which is 215. (61) I’ll try moving C above B first. (62) Going back to my previous and making moving B above E. (63) Nope. Didn’t do anything. (64) The next highest number is 139. (65) I’ve already moved F. (66) I’ve already moved A. (67) I’m moving D above A. (68) So I now moved D. And the schedule got worse. (69) Going back to my previous. (70) I can go to the next highest number which is 111 and move the jobs around there to see if it improves the schedule the total. (71) Nope so far. Oops. I already did this move. I just realized it. It’s too late. (72) OK. OK. It looks
like I moved most of these. (73) C,B... Yep, I moved these all so I’m going to now go into 2 position moves starting near the highest changeover cost which is 296. (74) Let me just jot this down. (75) I just I basically am drawing arrows so that I don’t repeat myself. (76) OK. I’ve got 6 moves. (77) I’ll start by moving D above F. Oops. (78) And I did not improve anything. (79) Next I’ll move A 2 jobs above C. (80) Definitely did not improve on that one. (81) Third will be moving F above B. (82) Hmm. I improved it. (83) I still have to go back to the previous because I have a few more moves to make on this. (84) I’m going to move B out to... (85) Did not improve it. (86) I’m going to move C below A. (87) Did not improve it. (88) And I’ve got one more move here to move F below D. (89) And I didn’t do any better so I’m going to restore the one schedule that I did make an improvement. (90) This is where I think I made a mistake, I’m not sure, on the ones we were given. (91) I don’t recall if I made a change in the schedule by moving two moves. (92) I remembered on my third one in that I should have gone back and made individual job moves, so I didn’t do that at all. (93) By the time I remembered it I didn’t make any changes with the double moves, so... OK. (94) So then I basically have to
start all over again. (95) I still have another high number, 2274. (96) So I’m going to start moving. (97) I’m going to start single movements again because I made a change. Again, per your instructions. I don’t know that I would have done this on my own, going back to the single, without your instructions. (98) And I improved it. (99) I’m going to go back to my previous, though, because I may be able to improve even better by another change. (100) Nope. Try one more here. (101) Nope. I’m going to restore the best achieved schedule, and that’s 6432, and I have a changeover cost of 175 which I’m going to go to next. (102) I’m going to make single moves again because I changed the schedule. (103) I’m not improving it. (104) Another one near the highest changeover cost. (105) And I’m not improving the schedule. (106) Hmm. Should have moved this one. I forgot this one. (107) Close, but my previous is still a little bit better. (108) OK. The next highest number, 93. (109) I’ve already moved G to A. (110) I’ve already moved A over D. (111) I’ve already moved D over C, so I’ve already covered that area. (112) Next highest is 89. (113) I could try moving F E. F above E. (114) That did not do any good. (115) That change did not do any
good. (116) Already did that already did that...
Looks like I've moved almost-- I've moved G,A,D. I've
moved C. I've moved B,F,E. (117) Then I would, now
I have to go back to the double moves. (118) Want me
to continue? Do I have to? No. OK. I have to. (119)
So 175 is the highest. (120) I'm going to double move
now. (121) See if I can improve this. I doubt it.
I'll be surprised if I can. (122) Oh. OK. I did
better that time. (123) Go back to previous though
and see if I can do even better. (124) OK. Oops.
OK. OK. I moved G. (125) I moved A. (126) Next I
want to move D. (127) Move D two up of B. (128) Got
worse. (129) Now I'm going to move down. (130)
Move B out. (131) I'm making 6 changes around the the
highest changeover. (132) I did find one though one
that is better but I'm still not going to do it until I
go through all 6. (133) One more to go on this one.
(134) I'm restoring the best which is 6388. (135)
That means I made another change making double moves
which means I have to now go back and do single. (136)
Are you sure you want me to continue? This is not fun
anymore.

EXPERIMENTER: Continue if...If you were doing this at
work would you continue?
I mean, I’d say there has to be a cutoff point where, I mean. What is this, how many jobs? 4, 5, 6, 7!

I know they wouldn’t expect you to go through all the iterations. I think we’ve put a lot of time and effort into this. I do want to…Let me analyze the schedule and see if I have…Ooh. Hang on a second. I can’t stop now. I’ve rearranged the whole thing and my job A is now 25 days late. I am not going in to see my boss with the finished schedule right now or I won’t be working on this any longer. OK. I have to move A. I cannot stop here. Hmmm. OK. I’d better. What’s going to happen when I move A? Because I’ve been asked to complete A on time I may not end up with the lowest grand total. It’s a tradeoff. I’m going to have to move A. I’m still 25 days… Wait a minute. Let me go back to the analyze screen. I’m 25 days late on job A. I have to move it up at least to this area here in order to have job A on time. I’m going to try to move A so that it’s on time. It’ll reduce my inventory holding charges because it’s pretty high right now being 52 days. Well, I’m going to try moving it here first. Well, I’m one day late. That’s not acceptable. I have to
make one more move. (160) I’m bringing up my totals even higher. (161) Well, I’m OK here but I’m not happy with the total costs I just created in this schedule, but I have job A on time. (162) I’ve reduced the costs from where I was originally, but I don’t know if this is the optimal schedule with A still being on time. (163) Only 5 days slack, so I can’t really trade it with any of the other ones.

(Tape turned over. May be gap.)

EXPERIMENTER: Where are you in your thinking right now?

(164) I’m thinking that I should not have gone through all the steps I just went through. (165) Because A has the longest duration so no matter all the things I’ve just done in rearranging the schedule it’s not going to put A where it should be. (166) I probably should have used a completely different thought process. (167) I’m just looking at to see the information on both screens again. (168) I probably should not have spent so much time on changeover costs because the objective with A having to be completed by a certain time frame should have probably focused on that and not even...tried to work with changeover costs in the other areas but nothing that would have affected
A. (169) So I shouldn’t have gone through all those iterations because I moved A back and forth throughout all of those. (170) I should have put A in its place that I had at the beginning and then attack changeover in the other areas if I could have improved the schedule and stopped at that point because I had 0 net effects so I concentrated I would have concentrated on tardiness which is delivering on time and inventory which means getting things out of the inventory you know faster and not incurring a cost. (171) I probably should have spent very little time on changeover. (172) I think I messed up there. (173) So if I can just go back to that original schedule that I had had come up with somehow. (174) No. That’s that’s my thought process now, but I realized it too late. (175) I don’t...I had... I had 0 net effects across the board and I just I should not have worked that hard on changeover. (176) I should not have gone through all the iterations I did on changeover. (177) There’s no way to go back to that stage, right? (178) OK. And now that I’ve gone beyond that and I’ve and I’ve worked on changeover I have to go back. (179) Job A has to be one of the top three, so I guess what I can do is see if I can improve anything at all by moving A. (180) Nope. I’m just moving job A and
making sure that it's completed on time but I have a couple of options and see if I can improve reduce the cost. (181) That's a little bit better. (182) Going back to analyze the schedule. (183) Looks like by putting A at where it's at now, which is first, I don't need to have it that early in terms of looking at just periods late because I was I was well OK a couple of jobs below where it's at now but I have reduced the total overall cost by putting it where it's at now so this is the better schedule so far. (184) Let me take a look at...Ooh. (185) Let me see if I can improve by going back to net effects. (186) I do have a large negative 216. (sneeze) Bless you. (187) Let me modify let me move these two jobs around. (188) I feel like I'm starting from scratch again but I had forgotten my main objective as I got carried away with the steps here. (189) Hmmm. I did better. (190) OK. Go back. I am improving it now. (191) Go back and analyze a few more of the jobs. (192) OK. Looks like I'm going to be changing a couple more here. (193) Should be OK here. OK. B above F. (194) It still doesn't affect A and I'm hopefully going to improve. (195) Nope. I didn't. I'm going to go back to previous. (196) Because there was another move I
wanted to make here. (197) I wanted to go back to the other one I had negative. (198) I think it was B and C. Which will still move B. (199) Yes, OK. We'll move B and C and see what happens. (200) I just moved B and F and it gave me higher cost for total. (201) Nope. This isn't any better. (202) I'm still going to have a problem wherever B is. (203) I have to do something with B. (204) Inventory holding charges 1148. (205) I'm not late with it now. (206) This screen doesn't tell me enough. (207) Second swap, B and C. (208) Not going to affect tardiness cost. (209) Kind of wish I would have stopped from the very beginning when I moved A to the very top and I had all positives and I should not have even looked into changeover costs again except for maybe within here without if I could have gotten a better number a lower cost number but there's a couple of jobs here that are out of place and I don't know which ones. (210) Duration doesn't help. (211) Inventory cost per period. (212) Let me see if there's another screen. (213) "H" High/Low. I didn't use this the other day. (214) No, that's not telling me anything. Strategy help. Identify change in schedule... Strategy help doesn't help because that's giving me your strategy for... How do I get out of here? End. (215) Actually,
my changeover costs are are low in comparison to where they were before. (216) They’re much lower. (217) My job A is on time. (218) Oh, my inventory holding costs are high for B. (219) I wonder if I switch B and E. (220) E has very low holding charges. (221) Oh, I can’t switch them. OK. (222) Let me try...

Nope. I don’t want to hold B any any longer, so I don’t want to move it in this direction because then I’ll increase my inventory holding charges. (223) I just tried to move it above E and that didn’t do me any good. (224) I think I’m going to end it at this point. (225) I tried B and F. (226) I tried B and C replacement. (227) Analyze the schedule once more. (228) B 12 days early. 41 Inventory cost. Duration 6. (229) I don’t understand why I can’t change these two and be OK because I wouldn’t be late so I wouldn’t be incurring a tardiness cost. (230) Am I looking at this right? (231) Try that. I thought I made that move. Just try it once more and see see if that... I think I did this one. (232) I just want to make sure because it looks like I would not incur that tardiness cost. (233) Nope. It doesn’t do any good. (234) I must be affecting...Oh, my changeover got really high. (235) There was the problem. (236) I brought the
changeover cost over here to like 200 and something so that must have been where although I kept the inventory and tardiness down I increased changeover. (237) I think I'm going to stop at this point because I don't see what else I can do. (238) Let me see one other thing. (239) See if my bottom ones I have any negative. (240) Nope. Good there. Good there. Anything with B and there's nothing I can do there. (241) I'm going to stop at this point. (242) Now. How sure are you that you found the lowest cost schedule. I'm not very sure but I think I got it down fairly well. I'll say a 3 and a 3.
TRANSCRIPT 4 -- EXTENSION

EXPERIMENTER: OK. Now on this one we’re going to change the situation slightly. For whatever schedule you come up with you have to make sure that job A is not late. So, you’re still trying to get a low cost schedule, low total cost,

(1) OK. Low total but...

EXPERIMENTER: but whatever your final solution is, job A is not allowed to be late.

OK. A not allowed to be late. (2) OK. So I’m writing it down so I won’t forget. (3) Are you able to go back if you get to that point? Because I know I forgot the two moves. (4) OK. Alright. So now the latest so now I’m going to go with the modify the schedule. (5) Analyze the schedule. (6) No. Yeah. OK I do the I&T net effect. (7) The best is -174.

(8) I’m going to modify. OK. (9) OK. Tardy is the second column. (10) OK. Now I’m going to use the tardy charge as my means of making my decision. (11) So 989 is the largest. (12) So I’m going to go one above 989 and move it up one. (13) OK. That’s higher so I’m going back to my best. (14) I’m looking at my totals for my tardy charge. (15) I’m going to go one job under and move it up one. (16) Which is lower so I’m going to make this my best. (17) And A...no job A
is the one that I should be concentrating on. (18) So instead of looking at the total I should be looking at job A tardy charge. (19) Let’s try moving job A up one position. (20) It was 400 last time. (21) Now it’s 256 which is lower. (22) Also I have to look over...but in the total is higher. (23) I’m going to go back because it’s higher on total. (24) Let’s try...the highest is 202 the one above moving that right there. (25) OK. Job A is 400. (26) Total make this the best. (27) Go one job under. (28) 92. Tardy charge. (29) No. Everything’s higher so go back. (30) Since we’re concentrating on A just out of curiosity I’m going to take A and move it up 2. (31) Made A smaller but in total the grand total is more. (32) I’m going to go back to my best and I’m going to take A since A has to be zero and move it up three. (33) OK. This is higher. (34) I’m going to go back to my best. (35) Previous was 5637. (36) I’m going to take A and I’m going to move it up four. (37) I have no strategy here. (38) I’m just trying this. (39) This is higher. (40) I’m going to go back to my best. (41) And I’m going to go back to the theory of using the largest number. (42) And I’m concerned with tardy charge so I’ll use my largest job and move it one position above up one position. (43)
That's higher. (44) I go back to my best. (45) The one under up one. (46) That's still higher. (47) Going back to my best. (48) OK. Maybe if I move by my changeover it will change my tardy charge, so I'm going to try that. (49) One under. (50) It's higher. (51) I'm going to go back to my best. (52) Two jobs under move it up one. (53) That's higher. (54) I'm going to go back to my best. (55) My next highest number move one position. (56) Oh, that wasn't the next highest number. (57) Back to my best. (58) 139 was my next highest number. (59) So up one. (60) This is still too high. (61) I'm losing control here. (62) Let's see. Move that up one. (63) Now, I'm going to move two positions at 139. (64) 111. I think I should write these down the numbers that I'm doing. (65) I think if I move the numbers at the...one down up. (66) This is the best take that one. (67) OK now you use since A has to be zero I'm going to try moving A up one. (68) That's better but my total is higher. (69) Let me try using this here and then going with my best. (70) I tried this one to bring my A down to zero. (71) A is coming down some. (72) I'm going to try moving A up one more. (73) OK. I have A down to 16. (74) I'm going to try moving A
up again one position. (75) I have A to zero. (76) I’m going to make this my best. (77) No, I just messed up. (78) I can’t make that my best. (79) How do I restore previous? To restore previous hit "P." OK. (80) Now I have A at zero. (81) Maybe if I change some of the jobs below A I can bring down my grand total. (82) So I’m going to move, let’s see, what’s high there? (83) My next highest number under A is 260 but I don’t want to move 260. (84) The C. Maybe I’ll try to move G between A and C. (85) Go back to previous. (86) Previous was lower.
(tape turned over)
(87) ...previous. Try to move B. (88) And the previous one is lower. (89) Now, restore best. (90) Restore previous. (91) I want to keep it here, but I don’t want to do best. (92) OK. How to...I don’t know. I’ll restore previous I guess. (93) A was at... Maybe if I move job A one. Let’s try it.
(94) OK. That’s higher so I’m going to go back to previous. (95) OK. Let’s try moving jobs under A directly under A down one. (96) OK. That’s higher.
(97) Go back to previous. (98) That’s lower so I don’t know how to keep this. (99) Let me see...operating help. (100) I don’t think I can undo a best best achieved. (101) I can change a best.
(beep) That won’t help. (102) OK. I want to
go...compute. (103) Restore previous. (104) Restore
best. OK. (105) I’m going to hit finished here
because I don’t think I can do anything else to keep A
as zero. (106) I’m sure. Please respond. How sure
are you that you found the lowest...I’m not sure due to
the constraints I had. I’m going to say I’m about a
three on there. (107) How sure are you my strategy
was...I’m not sure but I’m going to hit four because I
had A at zero. Are the values you entered correct?
Yes.
EXPERIMENTER: OK, now this next time through I want you to try and do the same thing which is create the lowest cost schedule, but this time you have a constraint which is that you have to make sure that job A isn’t late. So whatever you do, whatever schedule you come up with, make sure that job A isn’t late.

(1) OK. I’ll move A up until it’s not late. (2) Well, it’s not late. (3) I guess I’ll leave it there.

(4) I’m going to write down my negatives and do them all. (5) D is really tardy so I’ll try to move it back. (6) Ooh. Hmm. Close to quitting. (7) I don’t know. (8) D is really tardy. (9) You’d think you could move that back and save money. (10) Hmm. Well, I guess I’ll quit. (11) Oh, it’s late. (12) I moved those. (13) I’ve got to move those back.

(14) Yeah, that’s what I want to do. (beep)
TRANSCRIPT 6 -- EXTENSION

EXPERIMENTER: OK. Now on this next one we're going to change the rules just a little bit. Whatever schedule you eventually create, you have to make sure that job A is not late for that schedule.

OK.

EXPERIMENTER: So you're trying to come up with a low cost schedule but also one in which job A is not late.

(1) OK. (beep) Oops.

EXPERIMENTER: Don't forget to keep talking.

(2) Oh, I'm sorry. I'm sorry. I'm trying to find the lowest cost making sure that job A is not late so... (3) My best net effect was a -69 so... (4) Oh, darn. (beep) (beep) Oops. Some help...

(5) Because I forgot what I was supposed to switch around I'm going to start over again but that's probably still already in the computer. (6) I want to get the best. (7) I'm trying to eliminate this tardy charge and at the same time keep my costs low. (8) I wish I could talk more but I can't. (9) I'm not very good at this. (10) This is fun. It's like playing a game to me. (11) I could do this kind of stuff all night long trying to outwit myself. (12) Am I understanding this correctly? It doesn't matter...
what your total cost is as long as I don’t have this
tardy charge?

EXPERIMENTER: Well, you want to try and do both. You
want to try and keep a low cost and also not have the
tardy charge.

(13) OK. Hmm. Hmmm. Let’s see now... Hmm. This
strategy is not working at all.
(tape turned over)

(14) I am still plunging away trying to get the least
cost here. (15) Let me see. This is what I’m going to
do here. (16) Hmm. Take... I’m getting worse. (17)
(beep) Oops. My strategy is all gone. (18) Let me
see. I didn’t mean to do that.
(tape ends)
TRANSCRIPT 7 -- EXTENSION

EXPERIMENTER: OK, now the next one you do there’s a slight change in the rules. You want to come up with a schedule that’s the lowest cost schedule possible, but you also want to make sure that job A is not late in that schedule. So you have to work around that limitation that job A has got to be done on time.

(1) OK. Smith’s rule again. OK. Now this time, let’s see. (2) They’re all random but the random one. They’re all in random order but the random one. (3) Smith’s rule, shortest processing time. Let’s see, inventory cost, tardiness cost, changeover cost, and the total cost. OK. The inventory cost...Ooh. (4) These numbers fluctuate. (5) It seems like the inventory cost is very high where the tardiness cost is low on the two the weighted slackness and the least changeover. (6) Hmm. Changeover costs. Hmm. Shortest processing time was cheapest so we’re going to modify it. (7) Let’s see. Job A cannot be late so it’s got to be completed in 52 periods. Period completed 52? (8) I guess it’s got to be completed...I don’t know. (9) Let’s analyze it then. (10) Evaluate. (11) The best D and C and job A...Ooh. 25 periods late. 25 periods late. Wow. (12) D and C have to be changed definitely so modify them. (13) D goes where—wait, wait, wait, no. C

(turn tape over)

(18) Best. 84. D and G. (19) Modify D and G.
(20) Better. Analyze it. (21) Now, job A is still 25 periods late. (22) Fix it. (23) Evaluate. Evaluate. (24) The best is E and B. (25) Modify. Made it worse. Hmmm. Let’s see. (26) Go back and analyze it. Hmm. (27) OK. If I can’t improve on the net effect on the inventory and tardiness charges...best computed net effect can’t improve on that and job A is still 25 periods late and it can’t be. (28) The inventory cost is 34 dollars. (29) Oh, the job duration is 10. (30) No wonder. It takes the longest. (31) Tardiness cost per period 16. (32) Well, go to step two. See what we can do. (33) If it costs the most...400 dollars. (34) Make it the first job. Move it A underneath E. (35) There. Moved it two times towards the top. (36) Now. Hmm. Job A. Oh. Periods completed 19? Wait it took 25. (37) Wait a minute. Tardiness charge zero. Job A? What in the world... (38) Analyze the schedule now. (39) Hey, job A is negative eight periods late so it’s not late.
(40) Let’s see what the net effect is now. (41) Hey, job D and G. The best is 216. G and D so go to...G and D. (42) Take D. No wait. Which costs the most? G costs the most so I’ll take G and move it above one two above E. (43) See if that will work. No it didn’t do anything but let’s see what happened to A. (44) Analyze the schedule. (45) Oh, A is a period late. Shoot. (46) Evaluate. Hey, F and D. (47) 270 A and C, hah hah hah. A and C. (48) The best 270 -270 if we switch A and D. (49) Wait a minute, if we switch A and C the inventory and tardiness charge would have a -270 net effect. (50) Modify it. Oh, shoot, what was that? A and C. (51) Oh, shoot, what was that? God, how can I be so short...I lost it. It was A and C, A and C right? Yep, A and C. OK. (52) Modify it. A... (53) Let’s see which one do I want to move. Tardiness charges zero zero 84 16. (54) If I move C below D, D kind of skoots up a little bit. (55) Hmm. Let’s see. What did that do? 92. 92 is the lowest changeover charge. 92 123 120. Let’s see. (56) Analyze it. (57) A’s still late. (58) Go back to the beginning. (59) Oh, look. There’s one. (60) The best G and E. (61) OK modify G and E. (62) OK. Now I’m going to just stick by and take E and do
something with it. Take it the other way and put it above B C D. (63) Let's see what that did. Well, it got a higher grand total. (64) Analyze it. (65) Hmm. Oh, A is 11 periods early. That's good enough. (66) Modify it. OK. (67) The best achieved was 6666. (68) 7802. Hmm. How come it was...

(69) Let's see. A is not tardy. (70) In fact it's 11 periods early and the other ones are earlier yet. (71) -34,6,4,21,4 and 26. The cost... OK. Well, how come it costs more? (72)Hmm. Modification schedule. (73) Inventory holding charge, tardiness charge. Nothing nothing. (74) 250...Ooh, look at that 882 tardiness charge. (75) OK, let's move them around and try not to try not to disrupt A or we'll make it late again. (76) Let me see, B,A,G,F,E,D,C. OK, move F. (77) No. Take take take this tardiness charge...no, go back and analyze. (78) G and F both take 9 periods. So does C. But G doesn't cost that much. (79) G to F doesn't cost as much as D to C. (80) Work with C. (81) C. Take C and move it above E. Yeah, see how that works. (82) Get E down lower. (83) Oh, gosh. And I'll bet you A is late again. (84) No, it's still early. (85) OK modify it. No, wait. What did I do? Go back. Go back. (86) Take, now, let's see, B,A,G,F... Move C and E around. (87)
This is supposed to be two steps. Let's see. (88) I'm going to move E above... No, delete that. (89) We'll move D above C and then I'll move E down a little bit more. (90) No tardy. Oh, there's the tardiness charge. None for A. (91) I could have seen that before. (92) OK. None for B, none for A, 258 for G, 64, 57, 82... No no no. (93) Go back and analyze this thing again. (94) Alright evaluate. C and E that's the one. (95) Wait. Best. C and E. That's what I'm going to switch around. (96) Yep, I'm getting there. OK. (97) No, I don't want to take E up there. (98) That will mess up A and B. (99) C and E. Well, let's just switch them around. (100) Just the onezies, twozies. I'll just take them one at a time again.

(101) Hmm. That didn't help but it's a little lower. (102) Analyze that. (103) D and E now. D and E. That's what I'll do next. (104) D. Move D... switch D to E. (105) That's not going to save us any money. That's only 4 dollars. Let's see. D to E or E to D. Hmm. (106) Take D and move it above G just for the heck of it. (107) Just for the heck of it and see what happens. (108) Oops. That didn't look good. (109) Uuhh. Oh, wait a minute. Got less tardy charges. (110) B, A and D are not tardy. (111)
Analyze it. (112) Great. And they're early. (113) OK. Evaluate. E and F. F and E. 'OK. F and E. (114) Wow. We'll save aloit there. (115) Oh, look at this. (116) Best. F and E. (117) OK modify that. (118) F and try to move F down to here. (119) I'll take it two steps. (120) Hmm. Getting worse. (121) Take E and move it down to below F. (122) Hmm. Let's see what the best one looked like. (123) Restore best achieved schedule. (124) But that made A late. (125) Shoot. Darn it. A is late. (126) OK. Where was A before? (127) B and A. Modify it. (128) I'm just going to put A right here below B. (129) That's about the best. That's about the best. 7435. (130) Analyze it. (131) A's early 14. Oh, that's better yet. (132) OK. Modify it. I'm going to modify it. (133) I'm going to do one more modification. 7435. (134) Why should I do one more? That's good enough. (135) A is not late. It's 14 periods early. Let's see. (136) Period completed. 3,13,19... (137) Tardiness charges are zero, zero, zero. (138) Go back to the best achieved. Best achieved. (139) Now wait. Let me write this down. (140) E,A,B,C,F,G and D comes out to 7435 and if we go to the best achieved same thing only...E,B,C,F,G,D,A. (141) Change A with E. Oops. (142) One last try. Hmm. That's better. OK.
Let's see. (143) Restore...Analyze. (144) Hey, A is getting better. OK. (145) Now modify. A,E,B,C,F,G, and D. (146) It's A. It's E,B,C,F,G and D for 7319 and the best is 6666. (147) I don't like that C and F. Hmm. (148) Let me try one more. (149) Inventory holding charge, periods completed... No, no. No. (150) Previous. (151) Oh, wait a minute. That's OK. Forget it. (152) Finished. (153) Let's see, I'm positively sure that I got the absolute best. (154) How sure are you that the strategy you used...I don't know. I think I could have used a better strategy. Next.
TRANSCRIPT 8 — EXTENSION

EXPERIMENTER: (instructions not audible)

(1) Oh, gosh. I didn’t bring my notes with me. He went over this in class. (2) Use the shortest...isn’t that the one where you the...

(explanation)

(3) It’s the date due minus the...is that the one? The date due minus...

EXPERIMENTER: Right. Well, the date due isn’t displayed anywhere. Let me show you. See where it says tardy charge? The only way there’s a tardy charge on a job is if it’s late. So if you see a number there a positive number next to A it means it’s late. There’s one other way to tell if it’s late which is that if you ask to analyze the schedule, next to job A here it says how many periods it’s late. As long as this is a positive number it means it’s late. If it’s negative it means it’s early.

(4) Early. Right. I remember that.

EXPERIMENTER: So those are the two choices to tell whether job A is late or not. So this time when you create a schedule, just create it such that A isn’t late. You can make whatever changes you want but before you can finish you have to make sure that A is done on time.

(5) So you’re saying to create a schedule I can move around the like I was doing before?
EXPERIMENTER: Yeah. It's just that when you finish, when you have your final solution you have to make sure that job A isn't late. Right now it's late.

(6) But...

EXPERIMENTER: OK. Right now Job A is late because there's a tardy charge.

(7) Right. I mean I understand that. Right. (8) But I understand that but I thought these were already these numbers were given.

EXPERIMENTER: No. Those are computed.

(9) Oh, well, I don't know how to do that.

EXPERIMENTER: If A wasn't late you wouldn't have a tardy charge.

Right.

EXPERIMENTER: So if you put A earlier in the schedule then you will have less of a tardy charge. It will vary depending on where you put it.

(10) OK I understand that. May-- I don't think I know how to do that. (11) Let my try. Let me try because I don't think I know how to do it but maybe I know how to do it. (12) Now this is just trying to figure out how to make it so it won't be late. (13) Oh, OK. OK. Then I can move it around. OK. OK. I see. Let me go back to... (14) Do the best achieved
again so go back to that. (15) OK. Then the only way
I think I would probably do that is just by trial and
error. (16) I would all I’d try this 989 since that
would be the largest one and I would probably move G
maybe all the way up to the top. (17) But this would
be strictly by trial and error. (18) Although those
numbers they do appear to be...Oh, then I need to look
at the totals that’s what I needed to do. (19) That
might help just to let me know if I’m bringing my
totals down.

EXPERIMENTER: You want to bring the totals down...

(20) And I need to bring this down to zero.

EXPERIMENTER: No, you don’t have to bring everything
down to zero. Just job A. Other jobs can be tardy,
but only A...

(21) That’s right. You did say that. You know
what, you did say that. (22) And you’re right OK and
I forgot that you said that and I’m trying to make it
so that all the jobs aren’t late. (23) OK. OK. OK.
OK. Well, I think that then I think I would I would
take A out and put it in the first position and see
what would have happened here. (24) A. Well that’s
the...then that would would that be the answer?
EXPERIMENTER: Right. Well, it's one of the possible answers. A is not late, but...

(25) That's the main thing.

EXPERIMENTER: you still want to try and keep costs down. You're still trying to come up with a good low cost...

(26) But A not being late. Oh, OK. OK. (27) Well, I see that this grand total is much higher than the previous total so... (28) Now this would just strictly be by trial and error. (29) I would probably try to leave A in the first position and try... I would probably try I would manipulate the other variables to see what would happen. (30) I don't know. I would probably try to move this inventory holding charge out and move that all the way to the top. (31) OK. Now where is A. OK. A is not late. (32) That's even worse. I'm going back to my previous. (33) OK. Well, OK. What I would try to do now is try to bring my previous down a little bit lower. (34) And I would take my 95 and try to move that up at the top. (35) That's A's still not late but that's not the best answer. (36) I'd go back to my previous. (37) Maybe I would go down and move this 258. (38) I would put it behind the 95. (39) Have you had any bizarre reasonings? Anything really bizarre? (40) OK. Let
me see. Let me see. A’s not late. That’s good. And
that’s-- (41) No, I’d have to keep going back to my
previous total. (42) Well, OK well I’m not doing too
well with that so then I would probably try to
manipulate my changeover charges. (43) I would
probably... Try to take out the six and then bring it
up at the top. (44) Let me see. A is not late.
That’s OK and then...OK. (45) So then I see that this
is smaller than this so I’m getting somewhere so I’ll
stay with the changeover charges and try to manipulate
those a little bit to see if I can bring the number
down even a little bit lower. (46) Let me see... I
think I would try to take the 123 and move that out and
put that behind the 54. (47) Oh, OK. OK. Well, that
got a little worse so I’ll go back to the previous
schedule. (48) And then I take the 160 out and then I
put it... I put it right here. (49) Hmm. I’ll go back
to my previous total my previous schedule. (50) The
last one I did... Maybe I’d go down and take... I’ll
bring the 90 the 93 down and I’d bring it under the 54.
(51) A is not late. (52) No, I’d have to keep on
going back to my previous. (53) I’ll try maybe one
more thing on this. (54) I’m going to try one more
thing and I think I’ll bring the 160 all the way down.
(55) No no no no. I'll bring it down one more. (56) OK. A is not late and that's even worse. So I'll go back to my previous. (57) And then now I'm getting frustrated and disgusted so then I'll try then I'll try to work on the inventory holding charges for a little while. (58) Let's see if I can bring those down.

(59) I don't know. I think I'll pull the 246 out and put that up under the 125. (60) OK. Let me see. A there's no tardy charges on that. (61) No, that's even worse than before so go back to my previous total.

(62) I'll just try to move the 136 out. (63) When will you finish gathering all your data and doing your report? Whenevcr the report comes out, I'd be interested in finding out what your study concludes.

(64) Where am I? Oh, I got it down a little bit but I still...what did I do? OK. (65) I want to go back to the previous screen. (66) But you're doing, what's your total sample population? (67) Well, I'll just try to bring up the 125. (68) No, really and truly, when I was doing this Saturday and I kept I kept thinking the whole time that there must be some type of logical, methodical way to figure out the best answer with all these variables other than over than just sheer trial and error, the random method. (69) Because I think in a business organization they're not going to sit there
and play with this and do it by trial and error. (70) I was just thinking that we don’t we don’t know what the we may not know what the key is but I think in business they must use a particular formula, maybe something developed by the company based on their needs, based on their priorities. (71) But I know that nobody’s sitting out there doing this because I know it’s too time consuming. (72) Let me see...I’m going to go back to the previous screen. (73) OK. I think I’m going to try one more and I think I’ll move the 246. (74) OK. Now I’m at the point where I just I would go with bring the previous total back in and go with that go with the finished making my changes only because I’m frustrated and I don’t feel like I’m really making any progress. (75) I don’t feel like...I think that if I would come up with a number even lower than that then I think that might spur me to keep on going to try to, but all my numbers they keep coming up higher and I would probably use the one again for the same reason that I did before. (76) I know I’m not sure that I found the lowest. In fact I’m almost know that I had to have found the lowest cost... (77) And I would choose on a scale in terms of the strategy the best possible way...OK. I’d have to...Yes.
TRANSCRIPT 9 -- EXTENSION

EXPERIMENTER: OK. Now for the next one we’re going to slightly change the rules a little bit. You still want to try to get the lowest cost schedule, but you have to make sure that job A is not late. So whatever schedule you come up with, job A can’t be late.

(1) Yes, uh huh. Job A can’t be late. (2) Tardy charge...OK. (3) Job A being late. (4) I’m confused here how to determine that job A being late.

EXPERIMENTER: There’s a tardy charge next to job A.

(5) Uh huh. 402.

EXPERIMENTER: So that means it’s late. Also, on the analyze schedule there’s a column marked periods late. (rest not audible). If there’s a positive number next to A it means it’s late.

(6) OK. So I’ll have to use this net effect on this one. (7) This will force me to use this. (8) OK. The tardy charge or changeover time too. (9) OK. I didn’t know what that meant. (10) It’s different looking at it from that standpoint than to get the lowest cost. (11) OK. That can’t stay there. (12) OK. OK. Job A first then. Can’t be late. (13) OK. It’s not late anymore. (14) Best achieved. A little too high. (15) The more I work with this, the more familiar it is. (16) I understand it a little better
now. (17) Maybe I need to use that logic with some of these other ones. Looking for late... (18) This is the best achieved. (19) A was late. (20) Try the worst case for A and C. (21) Should be under the late charge there. (22) Still too high, though. (23) OK. Get this one out of here. (24) Gee whiz here. That one really didn’t like that. Hmm. (25) Bad move on that one. (26) Seems that wherever I move E everything goes high. (27) 74. Hmm. It needs to go to the top. (beep) Hmm. (28) 76. 75. Not too much difference. (29) Analyze the schedule again. (30) Try to get some correlation between E’s. E’s. 69. 210. C and F. (31) So I should stay with that. (32) What’s the next best? (33) That’s pretty good too. 210. So I’ve got... (34) Should bring me down. (35) Went up. (36) It’s not late but it went up. (37) Take A...Let me see. I should have kept A and C together. (38) Good net effect. (39) OK. A and C looks good on the net effect area so I’ll put D under A. Over A. (40) D over A had a good net effect also. (41) Over A, not E. Here. (42) Went up. D A C. (43) Why did that go up? (44) Because I had E at the top. (45) This should go down. (46) Not by that much. (47) I thought it would go down more. (48) C F G. (beep) (beep) Hmm. (49) Thought I had it figured out
but I don't. D A C F. 8000? Hmm. (50) I have a better one now. At least A's not late. (51) I need to go back and analyze schedule and find out what the...Try the high/low values this time. (52) That's a new one. I haven't tried this one. (53) Hmm. First last. (54) I suppose I should have looked at this one earlier. A F. A F. (55) These are good. (56) The last two aren't...wait. F and G. (beep) 63. F and G are the worst case. Let's see now that was...F and G. (57) C and G. C and G is the highest. OK. C and G is the highest. A is the highest...what's the lowest? (58) Try something different on this one. (59) Shortest processing time. (60) Try E at the bottom. (61) Gee whiz. It did not like that. (62) After G...OK. Try C over F. C over F. (63) 79? Gee, I made a difference on that one. (64) Gee, that really makes it...If I change G around or E those two seem to really make a difference. (65) Try it again just to make sure. I changed E with G. (66) It reversed quite a bit. OK. (67) Doesn't like that at all. Hmm. (68) G should stay at the bottom. (69) What else should go to the bottom? F. Try F. (70) Five. Nope. Let me go to my best schedule here. (71) F G A C...This one down next to F. 76. (72)
Right now it's in chronological order. (73) D is the best one. (74) Take it out of order and see what happens. (75) Hmm. It's defeating that. Hmm. F G... Change this one now. (beep) (76) Back up. OK. (tape turned over)
(77) ...high/lows. G and F. (78) OK. OK that was the best. G and F. Bring it up here. (79) Didn't like that. (80) Slightly different. (81) G. Doesn't like that. (82) Looks like that's the best I can do with this one.
TRANSCRIPT 10 -- EXTENSION

EXPERIMENTER: There's one other job set, but this time there's an added consideration. For political reasons, job A should not be late. So whatever you do in this schedule, you have to set it up so that job A isn't late.

(1) Isn't late?

EXPERIMENTER: Isn't late. Isn't tardy.

(2) Oh. I don't know how to do that, do I?

EXPERIMENTER: Well, just make sure that there's no tardiness cost for job A. So you can change the schedule as much as you like but in your final schedule job A can not be late.

(3) It can not be late. So you mean we can't have any tardy charge?

EXPERIMENTER: No, you can have tardy charges but not for that particular one. In other words, when you look at your schedule, next to job A there can not be any tardy charge.

(4) Oh, OK. I don't know if I understand this, but... OK. Let's see what happens. (5) Can not be... job A.

(new tape)

(6) OK now, I'm going to go over and analyze the schedules.
(7) Job A. I don’t understand that but...

**EXPERIMENTER:** You see job A? See where it says periods late in the last column? Right now it’s 25 periods late. In the final schedule, it can’t be late, so if you look at the periods late column it will either be 0 or a negative number. It’s not allowed to be positive.

(8) Oh, OK and changing them around is what’ll help me. OK. Alright, we’ll see.

**EXPERIMENTER:** The other jobs you can put wherever you like as long as job A is not late.

(9) Alright, OK, we’ll see. Job A can not be late.

OK. (10) Well, let’s see what the net effects are.

(11) This is added pressure here... on top of not knowing what I’m doing. (12) OK. I’m looking at F and G and C and F now for the net effects. (13) This is rough. (14) D and C negative D and C 174. (15) B and D. OK. And then E and B. Hmm. (16) Very interesting. OK, Let’s see here, I’m going to start with B. (17) I’m going to move B above E and make that the first job. (18) Hmm. 72. Let’s play with A down here. (19) I’ll make A the second job. (20) Did I want to do that? No, that’s worse. (21) I don’t want A up here. A. Let’s see, I want A... (22) OK. I’m going to put D... D above D above B.

(23) Oh, my gosh. That’s worse. Hmm, hmm, hmm. I
moved D back out. (24) OK. 75. OK. We'll put E back up at the top. (25) 73. I want D as the third job. (26) Hmm. 75 also, huh? I'll put C above D as the third job scheduled. (27) Hmm. So that's 68 so that's a little less. Now, D C...F. (28) F hmm above A. I think that's right. (29) That's 66 so that's a little better. Hmm. OK. (30) Let's go back over to the analyze the schedules. 66. (31) Oh, so A...is only late 16. (32) Now it was a negative. I see. (33) Let's see what the net effect of A and G is. -30. (34) Let me see how this is. Wait a minute. A, B, C, D, F, and G. So this is -30. (35) OK. Now F and A net effect -30. (36) D and F net effect, whoaoh, D and F is -228. (37) C and D the net effect is 228. (38) B and C the net effect -228. (39) And E and B the net effect -228. (40) Now let's go back to the modify the schedules. Hmm, hmm, hmm. (41) So I want E, B, C, hmm, E, B, C, D. Hmm. (42) I'll put this A above C and see what happens. (43) Hmm. That was higher. So take A and put it after C. (44) Hmm. It's getting worse. OK. I had a D with a negative 228. (45) I'm going to put that above the C. (46) Oh, no. That's worse. I'm going to take that D out and put the D after F. (47) It's worse. I want this F-- no, I want the A down as the next to the last one
right before the G.  (48)  Hmm. This is depressing here.  (49)  Did I do that? I want the F after the A.  
(50)  Oh. I’ll hit no.  (51)  Hmm. I quit. This is worse. E, B, C, D,.C, D. Put D before C.  (52)  Oh, man. That didn’t work. D I’m taking D back out and putting that after C.  (53)  68. So you’ve got E... now why can’t...that’s terrible.  (54)  I’m going to put A up here as the first job and watch it shoot sky high.  (55)  69. Hmm. I’m going to put A... Hmm. 69.  
(56)  Let’s see C I’ve got to get C back up here under B.  (57)  67. And this F has to go up here after C.  
(58)  That’s 65. OK. Now let’s go and see if we’re late.  (59)  It’s still late. Hmm. It’s late by 10. Ten.  (60)  OK. So we’ll go back to the modify schedules.  (61)  How could that be? 67. Hmm.  (62)  This is 8000 so that’s too high.  (63)  I put the G as the first job scheduled.  (64)  That blows that theory out the window.  (65)  So I’m going to take G and I’ll make it next to last.  (66)  Hmm, hmm, hmm. Hmm. 72. Then I’m going to make...So I’ve got E, B... I want to make C third again.  (67)  69. I want to see if I’m late. I’m still not.  (68)  Hell, I’m still late by one period. Hmm, hmm, hmm.  (69)  Hmm, let’s see what the net effects.  (70)  -216 G D. Hmm. F and G. A F.
C A.  B C.  Hmm.  E B.  Hmm.  What was this one?  C A.  
(71)  126 and 216 and A F was 52.  52 and 216.  Hmm, 
hmm, hmm.  (72)  OK.  I’m going to go back to the 
modify schedule and I’ll move A up...  (73)  Oh, no, no, 
no I don’t want that one.  (beep) Sorry about that.  I’m 
going to move A up to the third job.  (74)  And...  hmm. 
Have I had A up as the second job?  What the heck, 
we’ll go for broke.  (75)  72.  That’s 74.  Hmm.  I have 
to see take A out of there-- this is going to kill me-- 
and put it back after C and we’ll be back to some... 
(76)  6903.  OK.  So now we have E, B, C, A, F, G, D. 
That’s 6903.  Hmm, hmm, hmm.  OK.  (77)  Let me go and 
see how much late time I have.  (78)  It’s still one 
period late.  (79)  I can’t have any periods late. 
Damn.  (80)  Let’s go to high/low.  Let’s see what’s 
high/low this time.  (81)  OK, the lowest inventory 
holding cost was D and G and I have them at the end, D 
and G.  (82)  Cost...  The best was B and F.  Well, 
let’s see.  We’ll see here.  (83)  We’re going to go 
back to the modify schedule and try another theory 
here.  (84)  We’re going to go with B as the first job 
for...  (85)  6957.  And I’m going to put F as the 
second job.  (86)  69.  I’m going to put D as the third 
job.  (87)  Whoah, that’s too high.  I’m going to put A 
as the third job then.  (88)  Hmm.  76.  I’m getting
worse instead of better. What happens if I put E as the first job? (89) Hmm. 79 9. So I must still have a tardy charge, huh? (90) Let me go back to analyze the schedule. (91) I've still got that stupid one late period. Hmm. (92) How do you get rid of that garbage? (93) Let's get to the best achieved is 6555. (94) Now let's see how much late charge did I have there. (95) I had a 10. Ten periods late. Hmmmm 10 period 34. Hmm. (96) Let's go back to the modify screen and we're going to try... (97) If I put the A up here I'm going to put the A as the third job. (98) I think I did this before but... That's 70 92. (99) Well, let's see if I'm late. (100) Hmm. Now I have a negative. (101) Let's go back to the modify. (102) Hmm. Well that's dumb. I'll put F before the C. (103) 6941. Let's see if I'm late now. (104) Hmm. I have a -8. OK. (105) Let's see what the net effect is, I guess. (106) Hmm. D to G is 216 positive. (107) C to D is 178. (108) F to C 185. Hmm. (109) Hmm. OK. I'm going to move the A below the F and see what happens here. (110) Let's see. E, B...E, B...Hmm. Let me go back with this analyze schedules once more. (111) Oh, no, we can't do it that way. Now I have a positive. (112) Hmm. I'll
put this I'm going to put the I'm going to move the C
to before the G. Oops. (113) Let's see what happens.
Oh, wow. That's 70. 7099. (114) 6980. Hmm. A B... A B. I'm going to move the C to after the B. A B C.
(115) That's my 6 5. Now how late was I there. (116) I think that's the 10. Yeah, that's the 10. (117) OK. To tell you the truth here... (118) Wait a minute. We've got to go back and see the net effects
of this stuff. Did I do that? (119) D to G is 216.
(120) A to D is 40. (121) F to A is 40. (122) C to F 25. (123) E to C 25. Oops. (124) 69. Hmm. I
don't know. (125) I'll move the A-- that's my regular one-- after the C. (126) That's 67. Oh, am I still late? Oh, no. (127) I'm going to have to move the A.
(128) I'm going to move it to the E and see what
happens. (129) That's worse. Let's move it first and see what happens here. (130) 71. Now, let's see, how
late am I? (131) Oh, I'm early now. Hmm. (132) So
why don't I do this? Let's see. B, A, E, G, F, C, D.
(133) I'm recording A, E, B, C, F, D, G and that's
where I had 7157. OK. (134) I'm going to move the B
first and make that job one. (135) That's 7062 which
is better. 7062 B A E. (136) I'm going to move G up
after E. I'm going to have B A E G. (137) Hmm.
Well, we're going to go back and put in the previous
which was 7062. Hmm. (138) Well, let’s see if I put E
A I’m going to put E as the first job. (139) So
that’s 7092. The other was better. (140) So we’re
going to go back to the previous. (141) And, let’s
see, how late are we? Hmm. We’re early. (142) Now,
so we’ve got B A E we’ve got B A E C F D G and that’s
where we have... (143) I’m going to go back to the
modify screen... (144) 7062. OK. We’re going to go
back to the best achieved solution. (145) Let’s see,
and I’ve got 6903. I had E, E B, E B C and A was way
down there so they all... Hmm. B. (146) I’m going to
move B to the first job. Hmm. (147) That’s 66. Not
bad. So I got 6607 B (oh, hell, A is way down there) C
F A D G. (148) Now, let’s see, how late am I? (149)
10. Dang it. OK so now we have B E, B E. OK we’re
going to take A and move it up above the F. (150) 67.
That’s a little better. Let’s see how late we are.
(151) Oh, my gosh. Hmm. B A, B E, B E C D G. OK.
This A. (152) I know this is going to make I’m going
to put A as the second job. (153) Hmm. So I get 70.
That’s higher isn’t it? Yeah. I had 67 before. OK.
(154) What if I put E first? (155) Hmm. Still same
old garbage. OK. We’re going to go back to the
previous best. (156) 7062. Now, let’s see, how late
am I? (157) Hmm. I'm early. Oh, me, I'm getting
tired of this. (158) I've got the E up there. A B, E
B. OK. (159) We're going to take the A and move it
below the E so we'll make E the second job so it will
be B E and A. (160) 1762. That's higher. How much
time do I have? Hmm. There's time, though. Hmm.
(161) We're going to go back to the previous best
which is 7062. (162) We're going to move the A to
after the C. (163) I don't think this will work but
what the hell? (164) That's 6795. Now let's see.
(165) We're going to get to the analyze schedule and
we still have a stupid one in A so I can't do that.
(166) Hmm. So we're going to get to the previous. Can
I do that? Yeah. Let's see. (167) What should it be?
B and A. (168) I'm going to try one more thing and
then I'm going to quit. G. Move G up after A.
(169) That's 79.
(tape over.)
TRANSCRIPT 11 -- EXTENSION

EXPERIMENTER: For the next job set, there’s just a slight change in what I’m going to ask you to do. I still want you to try and find a schedule with a low total cost. The difference from what you’ve done before is that this time you have to make sure that job A is not late. So whatever other changes you make, you have to verify that before you say you’re finished that job A is on time. And you’re still trying to find a low cost solution given that constraint.

(1) OK. OK, first I’m going to analyze and try to eliminate all negatives net effects. (2) Go to the best. Let’s see. (3) Go to modify. (4) Analyze. (5) Best. D to F. Oh, enter. (6) Analyze. (7) Time to change D (?) and G. (8) OK. Analyze to see...I had a few more negatives. (9) Let’s see. We’re going to change B and E. (10) Analyze. (11) OK. Let’s go down just to make sure that there’s no negatives. Go machine. (12) OK. So far on this screen, A is on time. (13) Negatives mean late--mean early so we’re fine. (14) I’ll go back to modify and look at the changeover. (15) Let’s see. The highest changeover is going from C to F which is 296 so let’s play around with that one. Move C to F. (16) That’s not bad. What happens... Let’s see, 279. What if--I did C to F-- what if I did C down one more? (17) Nothing has happened to my A. (18) Oh, 65. The previous was better so I’ll go back to the previous
schedule. (19) Let’s see if I move F and E. Machine still moves too slow. (20) Nope. Still like the previous schedule. (21) OK. Let’s see, the changeover between D and A... (22) Nope, let’s work with B and E because that’s 229. (beep) Oops. Oh, press any key. Come on now. No such keys you could press. (23) Nope. Totals are too high. I don’t like that one so let’s get out of that. (24) Let’s get back to the previous. (25) Let’s work with D and A first before I start doing them two or three. (26) Ooh, that’s horrible. OK. I think I’m going to play around with F to C and move... (27) It’s a good thing the previous or best move is always there. (28) I don’t like that one. 6538’s better. (29) Let’s see if we move it up... Nope, delete that. Let’s see if we move... (30) Oh, enter. It’s not going to do it by itself. (31) Ah, I like that one better. 6521 and the A is still stable. OK. (32) What happens if we just change C and E? (33) No, I don’t like that. The costs increase. (34) OK, let’s see what happens if we play around with switching E... No, go back up and leave that. (35) Let’s go to F and G and see what happens. (36) No, I don’t like that. Go back to the previous. (37) OK. Let’s try F and E. (38) This becomes like
a puzzle, something where you want to break the code.  
(39) Enter. Come on, machine, I pressed enter once.  
(40) Oh, that’s horrible. (41) Let’s play around  
with what happens if we move A two spaces up. (42)  
No, I don’t like that. (43) What happens if I move G  
two down? (44) Nope. Hmm. That’s better. Let’s  
see. (45) My A period completed is still on time.  
(46) There’s alot of changeover cost between C and E.  
(47) My costs are down. Let’s see. Let’s check this  
to make sure I’ve got the best. (48) Nope, go back to  
previous. (49) Nope, go back to previous. That  
didn’t work. (50) Let’s move C. (51) Oh, I like  
that one even better. 6343. OK. (52) For some  
reason I feel like the changeover charges should be  
lower, I mean, not so many extremes like from 33 to  
229. (53) So let’s play around with B to E and just  
see what happens. (54) Nope. It’s low but it’s not  
the best, so let’s go back to the previous schedule.  
(55) What happens if we move B two places? Enter.  
(56) Nope. It got worse, tremendously worse.  
(57) I haven’t played around with E. Moved E to G.  
(58) Nope. Let’s move up... delete that. Come down  
to E and move it 1 2 3 there and see what happens.  
(59) No, that’s bad. Go back to previous schedule.  
(60) Hmm. See what happens between switching D to F.
(61) No. This is let's just double check F to A although I think I did that and now this is becoming...
Enter.  (62) No. Let's try F to see what happens.
(63) No, I don't like that one either. Let's see.
(64) A is on time and I think I feel like I've exhausted all possibilities.  (65) I'm going to play around with E to C.  E going to C...
(tape turned over)
(66) ...Golly. Finished making changes. Job...Press any key.  (67) "F." Job A is currently late. Please modify the schedule so that job A is not late.  Oh, OK.  
(68) Hmm. I'm going to go back to analyze so I can see...periods late 22.  (69) You were thinking backwards.  (70) Back to modify.  (71) Gosh, that's horrible. I don't like this cost.  (72) Let's just check to see if A...  A is still late.  (73) What if we took A... Delete that.  (74) Gosh, the costs are horrible. I don't like the cost.  (75) Let's play around with this. Are you completely finished making changes...I'm going to say no.  (76) I'm moving B and A.  (77) A is no longer late, but I don't like the total cost.  (78) That's worse so we'll go back to previous.  (79) I'm going to switch B and E.  (80) And A is better.  (81) Let's see if I switch... I
switched B and E. (82) Well, the previous was better. (83) Let’s move B and G. (84) No, previous was better. (85) What happens if we move B two? Enter. (86) Oh, that’s much better. I like that. (87) Let’s see what happens when we switch D F. Enter. (88) Oh, gosh. That’s horrendous. (89) I already did this. I’m switching A and E and I know it won’t work. I know it’s not my lowest. (90) Previous. Hmm. (91) I think this is the best I can achieve. Well, I’m not as confident, though. (92) Switch C... (93) Nope. Go back to previous. (94) OK. I think I’m going to quit. I’m going to finish this. (95) Are you completely finished making changes? Yes. (96) My confidence on this one is lower. I don’t feel as... It’s probably even lower than that. (97) I’m going to press no because I want to change this one. Four. Four. Yes.
TRANSCRIPT 12 -- EXTENSION

EXPERIMENTER: OK. There's one other job set, but this time, whatever you do, you have to make sure that job A is not late. So you're still trying to get the lowest schedule you can, but you're also trying to make sure that job A is not late.

Try to keep talking.

(1) OK. Alright. I've got to move A somewhere. (2) It's 25 periods late the way they've got it set up. (3) So I've got to get it up to zero or better. (4) I hate puzzles. I hate pyramids. (5) OK. Let's go through the same process, I guess. ...change the way to do this. (6) OK. G A is 30. (7) F G (Oh, hit the key, Gus) 63. (8) C F wonderful 181. (9) D C - 174. B D 210. E B -69. (10) So we have two to play with... A didn't get in there. (11) I'll play with B. (12) D. G... and should be these two. D C? OK. (13) There's ...slack... on A. (14) I'm going to try it the easy way first. (15) 73. 174 I want to play with D and C for a while. (16) Wrong direction, Gus. OK. Hmm. (17) Oh, alright. That's not the one I want to play with here. Wait a minute now. (18) OK. E B where are you? (19) No, I don't think that's what you wanted to do. (20) Where's B? Bingo. (21) E B OK. Let's go play with D C some more. (22) A alright. A
is still in there. OK. (23) OK G. Let’s see what...
(24) Hey, score one. (25) We still on time? Thank
you very much. (26) Alright. Let’s move this guy a
little bit and see. (27) I don’t think that’s a good
move. Wrong number. Damn it.
REFERENCES


