Fueling the Innovation Process:  
An Experiment and Field Study on the Relationship Between  
Explication and the Noticing of New Variables

by

David Gregory Rabkin

A.B. Engineering Sciences, Harvard University, 1983

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the
Massachusetts Institute of Technology
June, 1995

© David G. Rabkin, 1995. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part.

Signature of author: ____________________________
Sloan School of Management, June, 1995

Certified by: _________________________________
John S. Carroll, Professor of Behavioral and Policy Sciences
Thesis Supervisor

Accepted by: _________________________________
Birger Wernerfelt, Chairman, Ph.D. Committee
Sloan School of Management

ARCHIVES
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

-1-
JUN 14 1995
To Adrienne and Mitch
Fueling the Innovation Process:  
An Experiment and Field Study on the Relationship Between Explication and the Noticing of New Variables

by

David Gregory Rabkin

Submitted to the Sloan School of Management on June 5, 1995 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

Today's firms function in a world where technological innovation, regulation, new political regimes, and the creativity of their competition ensure that the set of variables critical to survival and profitability change continuously. Noticing these variables is critical to decision-making and adaptation throughout the firm. Yet even in hindsight, people have difficulty with noticing because of the way they filter and interpret information. This dissertation identifies explication – the process of making one's mental models explicit through language – as an important stimulus to the process by which individuals notice new variables. As defined here, noticing has been studied by very few researchers. While explication has received more attention, its potentially useful role in stimulating noticing has not been postulated.

The dissertation has two components. The first is a laboratory experiment demonstrating that explication is a catalyst to noticing new variables in a simulated marketing task. Further analysis is used to discover and test the set of motivational, script-based, and frame-based barriers to and enablers of noticing through which explication has its effect. The second component is a field study of a manufacturing organization. It uses a before/after design with an intervention. The field study develops the model further and explores influences on the explication-noticing relationship that were controlled or unobservable in the laboratory setting.

The findings have theoretical, methodological, and managerial implications. The dissertation highlights two constructs that have received little attention in the past, first hypothesizing then demonstrating a relationship between them. It proposes and demonstrates a series of mechanisms linking the two and, in the process, identifies two judgment errors that have not been addressed in the judgment and decision-making literature. It offers a new perspective in the ongoing debate on the validity of
verbal protocols and offers new reasons for concern about research that uses protocols to collect data. Finally, it offers a new agenda to managers concerned with their firms' ability to respond to changing environments. By identifying the mechanisms that drive noticing, it provides them with the information needed to develop well-informed, and therefore robust and effective, managerial practices.

Thesis Supervisor: John S. Carroll
Title: Professor of Behavioral and Policy Sciences
ACKNOWLEDGMENTS

My first thanks go to family: to my parents whose unshakable support has carried me through the most difficult of times; to my sister whose frequent reminders that she couldn't imagine getting a Ph.D. (and, by extension, that she thought I was nuts) reminded me that I was in fact doing something special; to my aunt who fed me countless dinners at Watertown's finest restaurants; and to my dear Patty whose timely arrival into my life and feisty spirit ensured my good cheer.

Special thanks go to John Carroll, whose intellect, common sense, and gentle humor contributed immeasurably to this work and to my experience at Sloan. I want to thank Stephan Schrader for his good advice, for the example he set, and for his friendship. John Sterman's broad perspective and red pen turned out to be great allies of mine in this effort, and I must thank him for his support. I am also grateful to Shmuel Ellis for his encouragement, insight and wit, delivered from halfway around the world via the Internet. Finally, I would like to thank Marcie Tyre who, from day one to the very end, was always willing to help and always supportive.

Two Masters' students, Gus Elmer and John Turbessi, contributed to my thinking very early in the process and helped keep me on the straight-and-narrow. Gus' expertise with macro programming and his help in implementing the experiment also proved invaluable. I need to thank my primary contact and friend at the field site; his enthusiasm was very important to me and his contributions critical to the field study. I am grateful to the many volunteers, both Master's students and workers at the field site, whose participation made this research possible.

Finally, I thank my fellow doctoral students for their friendship, support, and intellectual contributions.
A NOTE ON SOFTWARE AND DOCUMENTATION

The experimental task was programmed in the Microsoft Excel macro programming language (version 4.0 for the MacIntosh) and run on both MacIntosh and IBM personal computers. As noted in the body of the dissertation, the implementation of the experimental task represents work done jointly by the author and Augustus Elmer.

SPSS (version 6.1 for the MacIntosh) was used for all quantitative data analyses.

The data files, SPSS programs, and experimental software are available upon request for documentation purposes.
TABLE OF CONTENTS

1. Introduction ........................................................................................................ 15

2. Background and Theory ....................................................................................... 21
   2.1. Defining Noticing ......................................................................................... 23
       2.1.1. Mental Models ..................................................................................... 23
       2.1.2. Variables ............................................................................................. 23
       2.1.3. Adding Variables to Mental Models ..................................................... 25
       2.1.4. A Working Definition of Noticing ....................................................... 29
   2.2. Research on Noticing ..................................................................................... 30
       2.2.1. Noticing as a Construct ....................................................................... 30
       2.2.2. Noticing as a Dependent Variable ....................................................... 32
   2.3. Pilot Testing and the Emergence of Explication as a Variable ................. 38
       2.3.1. Goals at the Pilot Stage ...................................................................... 38
       2.3.2. The Birth of the Explication-Noticing Hypothesis ............................ 40
   2.4. Explication .................................................................................................... 42
       2.4.1. A Definition of Explication ................................................................. 42
       2.4.2. The Role of Explication in Cognition ................................................. 44
       2.4.3. Explication in the Context of an Organization ................................. 46
   2.5. Summary and Research Questions ............................................................... 47

3. The Link Between Explication and Noticing ..................................................... 49
   3.1. Overview ....................................................................................................... 49
   3.2. Experimental Design .................................................................................... 49
       3.2.1. Task ..................................................................................................... 49
       3.2.2. Experimental Treatment: Written Protocols .................................... 54
       3.2.3. Procedure ............................................................................................ 55
       3.2.4. Sample ................................................................................................ 57
       3.2.5. Variables and Measures ..................................................................... 58
   3.3. Results ........................................................................................................... 65
       3.3.1. Emergent Categories of Noticing and Recording Behavior ............ 65
       3.3.2. Descriptive Data ................................................................................ 72
       3.3.3. Noticing ............................................................................................... 75
       3.3.4. The Role of Explication ..................................................................... 76
       3.3.5. Sensitivity Analysis ............................................................................ 86
   3.4. Summary and Conclusions .......................................................................... 87
4. The Mechanisms that Drive the Explication-Noticing Relationship
4.1. Introduction
4.2. Methods
4.3. Barriers to and Enablers of Noticing
  4.3.1. The Motivation to Search
  4.3.2. Scripts that Affect the Search Process
  4.3.3. Frames that Affect the Direction of Search
  4.3.4. Discussion
4.4. Barriers/Enablers and the Explication-Noticing Relationship
  4.4.1. The Mediating Role of Barriers and Enablers
  4.4.2. Explication, Data-Recording, and Barriers/Enablers
  4.4.3. Discussion
4.5. Summary and Conclusions
5. Explication and Search Outside the Laboratory
  5.1. Introduction
    5.1.1. The Nature of the Experiment
    5.1.2. Goals of the Field Study
    5.1.3. Chapter Overview
  5.2. Methods
    5.2.1. Design
    5.2.2. Intervention
    5.2.3. Data Collection
  5.3. Results
    5.3.1. The Setting
    5.3.2. Noticing in the Observation Period
    5.3.3. Explication in the Observation Period
    5.3.4. Outcomes of the Intervention
  5.4. Discussion
    5.4.1. The Intervention and Its Two Levels of Meaning
    5.4.2. The Intervention and Barriers to Noticing
  5.5. Summary and Conclusions
    5.5.1. The Two Frameworks
    5.5.2. An Alternative Intervention
    5.5.3. A Note on Organizational Culture
6. Conclusion
  6.1. Recapitulation
  6.2. Discussion
  6.3. Summary
LIST OF FIGURES

3.1. User Interface for the Simulator ................................................................. 52
3.2. Deviation from Suggested Price, Coded by Advertising ............. 61
LIST OF TABLES

2.1. Variable State Transitions in the Modeling Process ..................26

3.1. Levels of Noticing and their Frequencies of Occurrence ..............66
3.2. Coefficient P-values in Decision-Predictor Regression Analysis ........................................................................................................67
3.3. Frequency of Agreement Across Measures of Noticing .................68
3.4. Bivariate Relationships between Variables .................................72
3.5. Data Overview ............................................................................73
3.6. Frequency of Noticing, Summarized by Treatment ......................75
3.7. Mean and Standard Deviation of Score and Time by Noticing .........76
3.8. Recording Behavior, Summarized by Treatment Group ...............76
3.9. Frequency of Noticing, Summarized by Explication ....................77
3.10. Logistic Regression Models Using Treatment and Explication to Predict Noticing .................................................................78
3.11. Frequency of Noticing, Summarized by Recording Behavior ...79
3.12. Logistic Regression Models Using Recording Behavior to Predict Noticing .................................................................80
3.13. Participants Reporting Time Pressure .........................................81
3.14. Time Spent in Minutes, Summarized by Recording Behavior ...82
3.15. Logistic Regression Models Using Recording Behavior and Time to Predict Noticing .................................................................83

4.1. Frequency of Barriers, Summarized by Noticing ......................94
4.2. Frequency Distribution of Barriers/Enablers ..........................141
4.3. Mean Number of Barriers, Summarized by Recording Behavior ..................................................................................................141
4.4. Logistic Regression Models Using Barriers and Explication to Predict Noticing .................................................................142
4.5. Frequency of Barriers, Summarized by Explication and by Data-Recording .................................................................144
4.6. Score, Summarized by Noticing and Data-Recording .......150
4.7. Remembering References to Advertising in the Case and Instructions .................................................................154

- 13 -
1. Introduction

Today's firms function in an environment where technological innovation, regulation, political regimes, social movements, resource constraints and the creativity of their competition ensure that the set of variables critical to firms' survival and profitability change continuously. When authors write about environmental uncertainty, they typically refer to uncertainty in the values of known variables: the supply of oil, whether or when high temperature super conducting materials will be developed to meet a given need, whether customers will continue to prefer neon colors. More than 10 years ago Schöhn wrote, "At least at the level of espoused theory, managers have become used to the instability of patterns of competition, economic context, consumer interests, sources of raw materials, attitudes of the labor force, and regulatory climate. And managers have become acutely aware that they are often confronted with unique situations..." (Schöhn, 1983: 239). Today, the situation is equally, or more, complex. The most challenging aspect of change, the factor that contributes most to the "uniqueness" to which Schöhn refers, is the introduction of important new variables that previously were non-existent, irrelevant or somehow held constant.

Researchers in the area of technological innovation often refer to the notion of the S-curve - a representation of the first-accelerating-then-diminishing returns of increased R&D effort for performance - to think about transitions from one technology to another (Foster, 1986). But the challenge from the innovator's point of view lies not in analyzing known performance variables but in identifying in the next generation of technology the new dimension(s) of performance that are presently unrecognized but will become the basis for future competition. In almost any domain of business, from the shop floor, where people learn to build new products or use new process technology, to the executive office, where people try to comprehend the new dimensions of the competition they face, the conceptual problem is the same: first noticing the new, relevant variables and then integrating them into a model that can help guide behavior.

Except in hindsight, these new variables rarely are obvious. Even with hindsight, noticing them is difficult because of the way we filter and interpret
information (Starbuck & Milliken, 1988; see also Lord & Foti, 1986). Yet they often can play critical roles in determining firm success. The firm that can notice them first may have an important advantage over others who are less aware or responsive. Since the people with the best opportunity to observe the many different aspects of the firm's environment and internal processes are distributed throughout the organization, wherever noticing is either implicitly or explicitly discouraged, the firm loses its best window onto some aspect of its environment or internal processes and thereby reduces its ability to notice and respond to change.

The actions taken to implement strategy or to respond to environmental change have consequences that cascade through the organization, affecting internal processes and thereby introducing new, important variables to be noticed. Those who have studied the process of implementing new production technology (e.g., Hayes & Wheelwright, 1984; Hayes, Wheelwright & Clark, 1988; Kazanjian & Drazin, 1986; Leonard-Barton, 1988; Tyre, 1991) note that its introduction starts a process of adaptation and problem-solving that may last for years. The problems that arise are often marked in their early stages by workplace symptoms that are new and unfamiliar. They represent a challenging and, when aggregated, important (see Hollander, 1965) issue for firms, highlighting the importance of noticing at all levels. Thus noticing functions both as the stimulus for innovation and as a driver of the problem-solving and adaptation processes critical to its success.

Despite its importance, noticing remains a poorly understood construct. From a theoretical perspective, noticing is relatively untouched territory. Typically it is viewed as an individual-level process based on aspiration levels (see literature reviews by Cowan, 1986 and Kiesler & Sproull, 1982) or as a group- or organization-level phenomenon based on interactions between people who share, perhaps negotiate, and otherwise enact their perspectives (e.g., Lyles & Mitroff, 1980). Both views assume that individuals have mental models of their environments and respond to shortfalls along specific dimensions. Neither questions how individuals learn that these dimensions (or variables, in my terminology) exist or are relevant in the first place. There are theorists who highlight the importance of noticing. Dawes & Corrigan (1974) stress the relative importance of identifying relevant variables.
compared to all other aspects of model generation (e.g., cue weights, functional forms, etc.). The small amount of empirical work addressing noticing directly attempts to describe the individual-level process and to demonstrate its importance (Klayman, 1988). This descriptive work, while insightful, does not suggest how one might intervene to manage the noticing process or to influence it positively.

The question that motivates this thesis is the relationship between the verbalization of mental models – an act I define more rigorously and then label "explication" – and noticing. In the past explication has been considered primarily as a means of collecting data (e.g., the debate on the validity of verbal protocols) or as a component of the communication process. Outside the field of psychotherapy, its role in stimulating the formation of new mental models has not been postulated. Driven by my interest in noticing, my sense of its importance, and observations from laboratory and field suggesting that explication may affect noticing, this thesis links the two constructs theoretically and empirically.

My approach has two elements, a laboratory experiment and a field study. Both studies are informed by the literature review and theory development in Chapter 2. This chapter also reviews the pilot work that led to the explication-noticing hypothesis.

The primary question, the relationship between explication and noticing, is expressed in Chapter 3 as an experiment that compares the behavior of people who do and do not explicate while performing a repeated decision-making task. It is a laboratory experiment, influenced heavily by work in the area of multiple cue probability learning (for example, Hammond, Summers & Deane, 1973; Todd & Hammond, 1965; and, most important, Klayman, 1988) and in experimental system dynamics (for example, Diehl & Sterman, 1995; Paich & Sterman, 1993; Sterman, 1989a, 1989b).

The analysis used in Chapter 3 is strictly quantitative. After establishing the link between explication and noticing in Chapter 3, I employ in the next chapter a qualitative analysis of the study participants' behavior as traced in real-time during the experiment and discussed in interviews with them. My goal was to develop a fuller understanding of the mechanisms through
which explication was having its effect. The chapter has two main sections: the first proposes a model of the induction process and then identifies factors that, by influencing the search processes of induction, act as barriers to and enablers of noticing; the second uses quantitative and qualitative analyses to demonstrate and further develop the relationships between explication, the barriers and enablers, and noticing.

The second study is a response to the tight control and relative sterility of the laboratory setting. Such a dust-free environment offers little insight into the problem of inducing explication in a real-world setting. The primary goal of the field study was exploratory, to identify variables that influence explication or that moderate the relationship between explication and noticing. If an effective implementation were possible, a secondary goal was to measure its effect.

The field study is presented in Chapter 5. I studied a group of production operators coping with complex and automated assembly equipment. The design called for two observation periods separated by an explication-inducing intervention. In the first period I studied the organization, recording day-to-day life in an effort to characterize the degree of noticing and explication on the part of the operators and to notice variables relevant to the success of an intervention. During this time, I worked with the production manager and supervisor to design an intervention to induce explication among the operators. In the second period, I implemented the intervention and observed its consequences.

Chapter 6, a review of the findings and discussion of their implications, concludes the dissertation. The two studies demonstrate a causal link between explication and noticing and develop a framework of variables that may play important roles in mediating and moderating their relationship. These variables are immediately useful in a practical sense, as grist for managers' thought processes; they also represent a series of hypotheses that will drive further work.

This dissertation makes several contributions. First, it studies the noticing of new variables, a construct that has been neglected for too long. It offers a description of the noticing process that is new, distinct from others before it,
and informative. It introduces explication, a variable that has appeared in literature but not linked to noticing, into a new domain and demonstrates its role empirically. It explores and demonstrates the mechanisms that explain the relationship between explication and noticing. In so doing, it introduces new theory and offers to managers new bases for taking action to help stimulate innovation in their firms. It adds new dimensions to the debate on verbal protocols, introducing new theoretical issues into a discussion focused primarily on methodology. It also offers new material for consideration as we design research and ponder our influence as researchers on the people we study.

Finally, and perhaps most importantly, this dissertation raises questions. There are many unanswered questions about the limits to which the experimental results in Chapter 3 can be generalized. There are specific hypotheses about judgment heuristics that emerge from the barriers identified in Chapter 4. These need to be developed further and tested. Stepping back a bit, I think that there is more inquiry to be made into enablers of noticing in particular, as well as barriers. As more are discovered, we may find that explication is neither the only nor the best way to influence them. Perhaps explication will prove to be the holy grail for those interested in influencing noticing; however, its most important role may prove to be the one it played here: as a key to identifying mechanisms that influence noticing. It is in the development of deeper understanding of these mechanisms and the ways in which they can be influenced that I see the greatest and most interesting opportunities for future work.
2. Background and Theory

As world events transpire, people within firms must cope with new technology, new sources of competition, new performance criteria for their products, the availability of new materials, and new forms of regulation. I have already pointed out that one challenging aspect of these changes is that new variables become important to firm success. The first, and a necessary, step in responding to these variables (as well as any others that have yet to be discovered) is a process labeled noticing, the focus of this research.

The importance of noticing is evident both at the practical level discussed so far and much more fundamentally. In presenting their model of the induction process, Holland, Holyoak, Nisbett and Thagard (1989) point out that "The success of a q-morphic\(^1\) model is fundamentally limited by the categories available for encoding the environment...The categories in the model are in turn dependent on the detectors that encode properties of the world" (Holland et al, 1989: 38). Noticing, which provides the material to be encoded is then an important, fundamental part of learning.

Unfortunately, human beings, and the institutions that we construct as our working environments, are not well adapted to meeting the challenge of noticing new variables. Since our thought processes tend to be schematic, our ability to perceive and recognize as relevant variables that do not fit our current thinking is limited (Lord & Foti, 1986). Arguably, our organizations have been built to cope with our individual needs for predictability and with the requirements for control of an earlier generation of management\(^2\). Today, of course, there is an effort to change our organizations. Researchers in the area of organizational learning, for example, are struggling to identify practices that can help people in organizations break out of their current, habitual ways of thinking and acting (e.g., Argyris, 1994; Cole, 1989; Senge, 1990). However, we are still less than perfectly adapted to our world's

---

1 They believe that mental models are structured as quasi-homomorphic models.

2 Juran (1964) defines control as the prevention of "bad change" and contrasts it to breakthrough, the creation of "good" (or necessary) changes.
changeable nature, and our organizations often do too little to compensate for our individual-level weaknesses and even reinforce them.

Noticing does have its costs. It may require effort and draw attention away from other activities, creating both direct and opportunity costs. If noticing occurs more rapidly than responses can be formulated, it can lead to lack of focus, instability of goals or plans, and failure to implement responses completely. Clearly, at the individual and organizational levels, the response to noticing must be managed under certain circumstances. However, I remain unconvinced that the process of noticing is so costly that it, too, should be constrained. Whether and under what circumstances noticing should be constrained and managed is both an interesting and open question.

However, I think it is safe to assert that generally there is a shortage of noticing in our organizations relative to the need and that noticing is critically important because it can provide competitive advantage to the firm that first notices new variables. My principal contact at the field site, a production manager, characterizes the issue as follows:

"If you need to get more efficient, more innovative, then you need to get at what's non-obvious. You can always get the obvious ones [variables]. To sit down and get the non-obvious stuff, [that] may save your butt. Looking around at the US, I'd say that most companies need to get at the non-obvious."³

Noticing functions as both the initiator of innovation and as a driving force supporting its successful implementation. It initiates innovation because it enables us to model our world, and perhaps our role in it, differently. It supports innovation by enabling adaptation, a process that should be framed as an active and creative one rather than one that is purely reactive and passive. Given its connection to organizational change and, as we shall see, our lack of knowledge about it, noticing is an important outcome for study.

³ Personal conversation, August 24, 1994. The requirement for confidentiality prevents me from revealing his or the firm's name.
2.1. Defining Noticing

Noticing is defined here as a part of the process by which people add new variables to their mental models. It is distinct from "noticing" a discrepancy between the actual value of a variable and one's aspiration level, the frequently used working definition of problem recognition (Kiesler & Sproull, 1982).

To clarify my new and more specific definition, however, I need to define the constructs "mental model" and "adding a new variable". Let me start by defining first the notions of mental model and variable. Then I can clarify what I mean by adding new variables. To do this, I will offer a model of the sequential states by which a variable is fully integrated into a mental model. Using this progression, I can then define noticing quite precisely.

2.1.1. Mental Models

Holland, Holyoak, Nisbett and Thagard (1989) define mental models to be built of if-then rules of two specific types. The first kind of rule maps the detected properties of the world (or, more likely some immediately relevant subset of the world) into states. These rules take the form:

IF property A [and/or property B, etc.] THEN the world is in state S.

The states are mental constructs used by the modeler ("cognitive systems" in the authors' parlance) to represent and understand the world. These rules combine to form a model's "categorization function."

The second type of rule, one or more of which comprise a model's "transition function", is used to predict behavior in the physical world by specifying transitions between states. These state transitions represent the effect of the passage of time, of actions that might be taken, or of both. Thus, mental models provide people with the ability: 1) to represent the world around them through categorizing its states and 2) to predict future states.

2.1.2. Variables

The terms used by Holland et al (1989) – detector, property, category, and state – refer either to the world outside the mind (e.g., they argue that "properties" actually exist out there) or to mental constructs (such as
"categories" which are clearly products of the mind). Unlike Brunswik's (1956) lens model, their terminology lacks a symmetry between the world and its mental representation. From my standpoint, this creates confusion.

Furthermore, they lack a term as general as the notion of a variable (defined below). Of their constructs, the "category" comes closest; but it carries specific implications that are not always appropriate and can become cumbersome at times. The authors' framework functions nicely when, in the first chapter, their examples consist of simple, categorical properties such as "uniform in color" or "striped." However by the end of the book they are serving up more challenging fare. For instance:

"The dispersion of category members over their dimensions of variation is not simply error variance in the usual sense; rather, it is itself a property [italics are mine] of the environment." (Holland et al, 1989: 185)

This particular property might be represented as a categorical variable (e.g., wide dispersion versus narrow dispersion). But why not have a construct less constraining than the notion of a category? Brehmer (1980) argues that human beings are both adept at and inclined toward using continuous variables and working with linear relationships. Given this, must we assume that the inputs to our transition functions are necessarily categorical?

Given the difficulty of framing the discussion that will follow in terms of categories, I offer the term "variable" as a label for our mental representations of the world's properties, and allow variables to be categorical or continuous. This definition should be quite consistent with the layman's definition of the term and most readers' likely interpretation.

Variables may be far more complex than mental representations of individual and simple properties, of course. Their values may be determined through examination of other variables, themselves potentially constructed from multiple variables. However, the top of such a hierarchical arrangement of variables can itself be termed a variable only if the modeler need not know the values of the component variables in order to make predictions. In short, the variable must be unidimensional, its values representable either as numbers or as a set of categories. If the modeler
cannot make predictions based on the variable's value but instead requires knowledge of its components, then it can no longer be regarded as a variable in the context of the model (it might be for some other model, however).

Variables, defined as our mental representations of the world's properties, are then the basic building blocks of our mental models.

2.1.3. Adding Variables to Mental Models

Adding a new variable to a mental model first involves establishing its role in the model's categorization function. In addition, if the variable does anything more than increase the accuracy of existing classification schemes – for example by serving as a new direct input to the transition function or by adding new categories to the potential values of an existing variable – it will necessarily change the transition function as well. Let me expand on and clarify these points.

In order to define noticing in a robust manner, we need to examine the events that lead to a variable's inclusion in a mental model. To do so, I define a series of steps through which variables progress before they are established as components of mental models. Each step corresponds to the variable's membership in one of a series of hierarchical sets of variables. Although several of these state transitions may be accomplished simultaneously in a single cognitive act, none may be skipped; a variable's membership in a higher-level set necessarily implies its membership in all the lower-level sets that precede. These sets and the progression are shown in Table 2.1 and described in the subsections that follow.

---

4 Holland et al (1989) use this term to label the way in which we map properties into categories. Although the word "categorization" may appear misleading (since in my thinking this function maps properties into variables which are not necessarily categorical) I choose not to relabel it. The construct appears too rarely in this document to warrant introduction of a new term.
<table>
<thead>
<tr>
<th>Set Name</th>
<th>Description of content</th>
<th>Status of variable</th>
<th>How a variable enters the set (progression)</th>
<th>How a variable leaves the set (regression)</th>
<th>Role of variable in the context of the specific problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universe</td>
<td>Universe of stimuli as yet uncharacterized as variables</td>
<td>Undistinguished</td>
<td>In the beginning...</td>
<td></td>
<td>Non-existent</td>
</tr>
<tr>
<td>Known variables</td>
<td>Universe of known variables</td>
<td>Distinguished, can be represented symbolically, not part of any problem frame</td>
<td>Distinguishing</td>
<td>Forgetting, reframing</td>
<td>Unused</td>
</tr>
<tr>
<td>Candidate set</td>
<td>Variables framed as potentially relevant to the problem</td>
<td>Framed as generally relevant (but not included in a specific hypothesis)</td>
<td>Framing a candidate set for a problem, directing attention</td>
<td>Framing out, reframing, forgetting</td>
<td>Building blocks for hypothesis generation</td>
</tr>
<tr>
<td>Tentative model</td>
<td>Variables with specific hypothesized roles</td>
<td>Hypothesized specifically as related to other variable(s) in the model</td>
<td>Hypothesis generation</td>
<td>Disconfirmation, reframing, forgetting</td>
<td>Testing, tentative use</td>
</tr>
<tr>
<td>Beliefs/Confirmed model</td>
<td>Variables believed to be causal</td>
<td>Believed specifically related to other variable(s) in the model</td>
<td>Hypothesis testing</td>
<td>Reframing, forgetting</td>
<td>Use, perhaps automatically</td>
</tr>
</tbody>
</table>
1. *The initial condition: undistinguished variables*

Initially, variables are undistinguished. Since variables are mental constructs, they must be created by a person in order to come into existence. The stimuli that they will represent, of course, do exist since they are properties of the world. But until they are mapped into variables they cannot be reflected in mental models or influence behavior. The set of undistinguished variables, then, is the infinite set of all variables that one might imagine to represent the universe.

2. *Distinguishing: How variables enter the set of known variables*

The first step on a variable's road to inclusion in a mental model is when it is "distinguished" from its background. At that point, a mental representation of some property of the world is constructed. Distinguishing a variable may be the result of identifying some new stimulus, or in the terminology of Holland *et al* (1989), a new detector. Alternatively, it may involve putting together already-known variables in new ways. Once distinguished, the variable can be named and represented symbolically. We now have a "known variable".

Variables can leave the set of known variables in either of two ways. At any point, the modeler may forget a variable, which then simply falls back into the realm of the undistinguished. Variables may also be framed out of existence if the same characteristics of the universe are distinguished into different variables; this kind of reframing may render a variable conceptually obsolete (as might be the case with X and Y when switching from rectangular to cylindrical coordinates). When such reframing occurs, then the old variable, if forgotten, leaves the set of known variables. If not forgotten, it may remain there but not progress further in the process.

3. *Hypothesizing relevance: How variables enter a model's candidate set*

The next step in the process is recognizing that the variable is potentially relevant to a particular model. The notion of relevance allows a variable to be linked to a model when its relationship with the model's other variables is not at all understood; relevance can take the form of the barest hunch. Nonetheless, a variable's transition from the set of known variables to the candidate set is a significant conceptual leap for the modeler. A variable's
entrance into a model's candidate set corresponds to formation of the following hypothesis: "variable X is somehow related to the other variables in this model"; the hypothesis does not need to state anything specific about the variable's role.

Variables can leave the candidate set by being forgotten or through reframing (both discussed above); in reframing, if the old variable is not forgotten, it leaves the candidate set but remains in the set of known variables. There is a third exit-path – to which I shall refer as "variable disconfirmation" – that also leaves the variable as part of the known set but no longer a candidate. Such a variable is assumed to be irrelevant to the model. I will discuss this particular inference at some length in Chapter 4. For the time being, just note that the hypothesis "variable X is relevant to this model" cannot be disconfirmed, first because no hypothesis can be positively disconfirmed but more importantly because it is an untestable hypothesis. Variable disconfirmation then is a matter of problem framing. One of this study's contributions is its distinguishing of this judgment and identification of its importance. As we shall see, variable disconfirmation can occur in several different ways, none based on strict rationality.

4. Hypothesis generation: How variables enter the tentative model

The modeler must hypothesize a role for the variable for it to enter the model itself; it must be specifically linked causally to some other variable(s) in the model. Most likely the variable will be introduced tentatively, as part of a hypothesis about its role. The degree of specificity may vary, of course. At a minimum, such a hypothesis must include a relationship between two specific variables. A fully specified hypothesis must indicate the values that both variables can take on and exactly how those values relate to each other.

Hypothesis testing alters the modeler's confidence in hypotheses. Should the modeler interpret the results of a test as disconfirming, the hypothesis will be dropped, removing it and (unless it is simultaneously replaced with a new hypothesis) the variable from the tentative model. Most likely (but not always, as we shall see in Chapter 4) the variable will remain in the candidate set, available for additional hypothesis generation. Hypotheses can also be forgotten, another action that leads to removal of the variable from the
tentative model. Finally, reframing may eliminate the variable entirely and therefore remove it from the tentative model.

5. Hypothesis testing/confirmation: How variables and their roles become beliefs

Over time and with experience, confidence in the variable's role may grow until its role has been "confirmed." I propose that confirmation, or belief as I prefer to call it, results in the unquestioning and perhaps automatic use of the variable in its hypothesized role. Note, of course, that consistency of use does remain an issue even in the presence of belief (Hammond, Todd, Wilkins & Mitchell, 1966). Sometimes, hypotheses containing new variables can enter the realm of belief immediately, bypassing anything that resembles a scientific notion of hypothesis testing.

Beliefs, of course, can be shaken. Although people exhibit great overconfidence in belief formation and a surprising ability to ignore information that disconfirms beliefs (Ross, Lepper & Hubbard, 1975), variables believed to be part of a model may be shaken loose if events can create doubt about them, returning them to a tentative status or, in the event of either reframing or forgetting, removing them from the candidate set altogether.

2.1.4. A Working Definition of Noticing

For the purpose of this study, a variable is considered noticed if it becomes part of the candidate set for a particular model. Noticing does not take place absent a model; a variable that merely enters the known set is not considered noticed. When a variable is noticed, its specific relationship to the goal or outcome of the model need not have been fully established; however, the variable must be immediately available as a candidate for use by the hypotheses generation process.

This definition implies that a person might be quite aware of a variable, as a known variable or as an element of some other mental model, and yet not have noticed it in a given context. It also implies that a variable may be "unnoticed" if a once-noticed variable is forgotten, framed as irrelevant, or rendered obsolete by an entirely new framing of the problem and/or its candidate variables.
2.2. Research on Noticing

2.2.1. Noticing as a Construct

The closest analog to noticing that appears in the literature is Klayman's (1988) notion of "cue discovery." He defines cue discovery as "identifying a set of valid predictive cues." His definition includes the processes of: discovering cues, incorporating them into models, and eliminating invalid ones. His study suggests first that people use outcome feedback to accomplish cue discovery, second that the presence of random error (as opposed to unexplained variance from as-yet undiscovered cues) slows but does not halt the discovery process or terminate it prematurely, and third that having the ability to design experiments\(^5\) helps the process.

With the exception of Klayman, very few researchers have actually required their subjects to distinguish variables. Bruner's famous work on concept attainment (Bruner, Goodnow & Austin, 1956) aims specifically to avoid the issue of distinguishing. In contrasting his research to Hull's (1920), he notes that:

"In studies inspired by Hull's procedure, it was not made clear to the subjects what it was about the instances presented to them that might be relevant... the number of attributes that a subject might consider as possibly relevant are close to limitless... So long as the experimenter does not know to which and to how many component attributes the subject is attending, it is impossible to control or understand the amount of information being presented... the process of how concepts are attained, given the abstraction of attributes, is greatly obscured." (Bruner et al, 1056, pp. 135-6)

In Bruner's studies, subjects were supplied a list of variables from which to assemble concepts; they selected and used, but did not distinguish, variables.

\(^5\) He matches experimenters, subjects who are permitted to choose the trials they use in a computer-based game, with observers, subjects who watch the experimenters' trials and outcomes and thus, Klayman argues, should be able to receive the same information. Of course, one can question if they really receive the same information given that the observers cannot communicate with their matched experimenters and therefore may be unaware of the experimenters' hypotheses.
Chapter 2

Of researchers who have required their subjects to distinguish variables, only Klayman (1988) has really attacked noticing as a construct. Hull (1920) observed that when subjects proceeded at their own pace, they learned to distinguish categories more rapidly when simple examples were presented first and followed by more complex ones rather than in the reverse sequence. In naturalistic situations, however, one cannot control the sequence or timing of examples. Thus the finding offers more to teachers trying to develop in others an understanding of a concept than it does to people trying to make sense of a 'world over which they have little or no control. Wason (1960), and those following in his tradition (e.g., Einhorn & Hogarth, 1978; Klayman & Ha, 1989; Mahoney & DeMonbreun, 1977), focus on hypothesis testing strategy and demonstrate the common presence in their subjects of what they term a confirmatory test strategy. This research is related to the topic of noticing because the confirmation bias minimizes disconfirming feedback and thus obscures the inadequacy of current thinking; if people view their thinking as adequate, are they likely to search for or even be open to recognizing new variables?

Much work appears to focus on noticing. For example, Einhorn and Hogarth write:

"Why are certain variables chosen as 'relevant' to the phenomenon in question while others are considered irrelevant or of lesser importance? This issue goes to the heart of all prediction problems although it has not received the attention it deserves." (Einhorn and Hogarth, 1982: 24)

Close examination of their argument reveals that they are concerned with how people infer causality among sets of known variables. The problem of how variables become known and enter the set of candidates for causality is not addressed. Similarly, the literature on individual-level problem recognition (see reviews by Kiesler & Sproull, 1982, and Cowan, 1986) presents a model based on aspiration levels along established dimensions. The issue of how those dimensions are established is not clarified. Those who do try to describe where the set of variables comes from, such as Pounds (1969), resort to people's past experiences, communications from others, and analogy as sources. Analogy has the potential to drive noticing because it suggests roles in a model that need to be populated by variables (see Senge, 1990, for his
discussion of the role of "archetypes"). Absent an analogy, however, this work only wards off the question temporarily. These answers prompt the questions: how did the person first distinguish the variable and how did the person communicating the model distinguish it? The issue of distinguishing new variables at its core remains unaddressed.

2.2.2. Noticing as a Dependent Variable

While very few people have studied noticing as it is defined here, many researchers have studied dependent variables that include noticing as a component. This work spans topics including creativity, induction, learning, problem formulation, organizational learning, and reframing. I group this work around a series of questions that they help answer.

Perceived need to notice and search

The inability to cope with a situation (about which one cared) presumably would lead to a perceived need to construct a better mental model. This might motivate search processes that could drive noticing. For example, one of Klayman’s (1988) principal and most interesting findings is that the presence of random error, in contrast to variance of equal magnitude caused by "hidden" cues, is a stimulus to structured experimentation. Thus, although the learning task in the presence of random error is more difficult, people learning Klayman’s experimental task did equally well under both conditions. I propose that the study participants experiencing the random error experienced greater incongruity; they were surprised more often. When people experience different results in what should be identical trials, cognitive alarm bells go off. They recognize that they are missing some critical piece of information and their perceived need to notice is raised considerably. Klayman's research suggests that the perceived need to notice leads to search which in turn leads to noticing.

---

6 These are cues that would be very difficult to discover. In his experiment, these cues were never discovered. Variance from hidden cues means that if when subjects try a given action more than once, they get the identical result. Where variance comes from random error, identical actions can have different results.

7 The two of us have discussed this and he agrees.
Chapter 2

Essentially any kind of disconfirming feedback might have this effect. Although I have already distinguished between aspiration-based notions of problem recognition (e.g., Cowan, 1986; Kiesler & Sproull, 1982) and noticing, problem recognition certainly can provide feedback that can lead to search and noticing. The perception of discrepancy is an important element of the induction model presented by Holland et al (1989). Essentially all learning, the authors argue, is driven by the failure of rules in one's mental models as demonstrated by disconfirming outcome feedback.

However, even when people receive disconfirming feedback, they tend to respond inappropriately from a strict probabilistic viewpoint. One characteristic of human information processing is conservatism. Edwards (1982) notes that, in the face of new information, people do adjust their opinions about probabilities and they do so in the correct direction; however, they also adjust them far less than they should given Bayes' theorem. People exhibit overconfidence in their ability to predict, adjusting confidence upward over time despite the fact that they receive no feedback indicating improved predictive accuracy (Oskamp, 1982). The actual feedback should disconfirm their high levels of confidence; but it does not. In short, our judgment heuristics make it less likely that any given level of disconfirming feedback will cause us to question our thinking and thus feel a need to improve our models.

It should be recognized that dissatisfaction with a mental model does not necessarily translate into perceived need to notice. The literature does not answer the logical question: Do people who are dissatisfied with their mental models go about improving them in different ways? Do they exhibit characteristic preferences between: the search for new variables; fine-tuning the coefficients associated with the different variables in the model; experimenting with functional forms; and trying out different combination rules? Any bias away from the search for new variables would moderate the perceived need for improving the model and the perceived need to notice.

Problem framing

Schema theorists have demonstrated that our mental models act as filters in perception, interpretation, and memory processes (Cohen, 1981; Lord & Foti, 1986). Surprisingly, they continue to have their effect as filters even
when they have been disconfirmed (Ross et al., 1975). Thus the way we frame problems can restrict the variables to which we will respond and, in particular, may tend to keep us from noticing new ones. If our frames become institutionalized as measurement systems, automated systems for data acquisition, and databases, then our natural inflexibility becomes further reinforced by the presence of infrastructure.

Schemas not only affect automatic cognitive processes, they also can bias more conscious search processes. In several different domains, research has documented a human tendency toward search patterns that confirm current hypotheses. Snyder and Swann (1978) demonstrate that when people are asked to determine through interviews if another person exhibits some characteristic (e.g., extroversion), they tend to ask questions that make disconfirmation unlikely. The odds of being labeled an extrovert by an interviewer determining if you are an extrovert is greater than that of being labeled an introvert and equal to that of being labeled an introvert by someone determining if you are an introvert. Significantly, this is not just a judgment error on the interviewers' parts; the interviewers actually ask different questions with the two framings. This biased search strategy appears elsewhere as well. Wason (1960), Klayman and Ha (1988), and many others have noted that people's current hypothesis about a situation affects the nature of the experiments they choose to make. Universally, this work demonstrates the confirmatory test bias and suggests that our heuristics for directing search help us avoid disconfirming feedback and thus prevent us from sensing the need to develop better models.

There is another type of framing that plays an important role in determining whether we notice and perhaps consciously search for new variables. Schrader, Riggs and Smith (1993) note that in real-world situations it is often unclear whether one's mental model can be improved. Furthermore, whether one needs to do so is a matter of framing, a judgment about model adequacy, rather than accuracy. Two people may frame a situation quite differently in terms of the adequacy of their models or the opportunities to improve them and thus may exhibit very different perceived needs to notice.
Cognitive modes

Some theorists argue that the effects of our mental models on noticing may vary on the basis of our "mode" of information processing (e.g., Chanowitz & Langer, 1980; Louis & Sutton, 1991; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977; Schön, 1983). Humans beings have two distinct information processing modes. In the first, we "run" well-learned scripts and examine only variables included in the script at the points in time indicated by the script. Noticing new variables during this automatic processing mode is very unlikely. Under these circumstances, our mental models act as constraints on noticing, as described above.

However, we are also capable of a second "controlled" mode in which we are open to new or discrepant information at any time. Louis and Sutton (1991) argue that people switch from an automatic to a conscious mode when they encounter novel situations or discrepancies with their expectations; or they may do so deliberately. Switching from one mode to the other appropriately is an important determinant of performance (Louis & Sutton, 1991). Use of an automatic mode when a controlled mode is appropriate will lead to insensitivity to non-scripted events and variables; use of a controlled mode when it is not required may take attention away from other important tasks and presumably would represent wasted cognitive effort.

Unfortunately, the presence of useful variables to be noticed will not lead to any perception of novelty or discrepancy. Only a perceived failure of one's current mental model would do that; in the absence of such feedback, then, only a deliberate choice to use a controlled mode of processing could influence noticing. Such choices, however, can be made, as practitioners of TQM and structured problem solving techniques make clear (e.g., Shiba, Graham & Walden, 1993).

Characteristics of the system being modeled

In discussing how people determine causality, Einhorn and Hogarth (1982) note a series of cues that can help people infer causal relationships between variables (and, I would argue, distinguish causal variables, as well). These include: temporal sequence – causes precede effects in time; similarity – causes that bear some resemblance to their effect are more likely to be
recognized; constant conjunction – the degree to which the variables occur together; contiguity in time and space – which make constant conjunction easier to recognize. These cues are all characteristics of the situation and cannot be manipulated. The authors discuss other cues; however they are primarily a function of the person doing the noticing and are discussed below.

I have already discussed the effects of random error in a system on noticing (Klayman, 1988). However, there are other characteristics of systems that may play roles, as well. Sterman (1994)\textsuperscript{8} observes that time delays make learning more difficult, as does the presence of feedback loops particularly in the presence of interactions.

**Characteristics of the individual**

In discussing cues to causality, Einhorn and Hogarth (1982) mention the use of analogy to suggest causal relationships. In her work on creativity, arguably an outcome to which noticing contributes, Amabile (1988) mentions the individual's "creativity-relevant skills", describing them as a cognitive style favorable to taking new perspectives on problems and application of heuristics for the exploration of new cognitive pathways. The mechanisms that March, Sproull and Tamuz (1991) suggest to maximize learning from rare situations (or even those that do not happen) also constitute such heuristics. One such mechanism that reappears in all this literature is the use of analogy. Analogy can suggest variables directly, or it can identify roles that need to be populated with variables. Senge (1990) advocates the use of "system archetypes" as abstract analogies to help people understand the systems in which they live and work. Archetypes suggest the structure of commonly found systems; through mapping the generic structure of the archetype onto their perceptions of a situation, people may be directed to search for and notice more variables.

Amabile (1988) notes additional characteristics of the individual that are associated with creativity, including: persistence, curiosity and energy, high degree of intrinsic motivation (i.e., interest in the task being undertaken), and

\textsuperscript{8} This paper nicely summarizes and integrates a series of his own experiments (e.g., Diehl & Sterman, 1993; Paich & Sterman, 1993; Sterman, 1989a, 1989b)
a willingness to take risks. While some of these features are probably immutable characteristics associated with the individual, it might be possible to influence others, such as the use of analogy or the willingness to take risks.

Organizational context

Amabile's discussion of the organizational circumstances supportive of individual creativity suggests how organizations may influence creativity. In her (1988) study, creativity was associated with the presence of organizational mechanisms for exploring new ideas, autonomy, the availability of resources including time, explicit encouragement and recognition of creative effort (not only results), a climate where failure is not "fatal" and where criticism and evaluation are not too harsh, and the presence of synergy between people rather than competition and defensiveness. She demonstrates that extrinsic rewards for creativity (e.g., money, promotions, etc.) actually have a negative effect because they undermine any intrinsic interest that people may have (Amabile, 1979). She argues that the organizational conditions listed here influence some of the individual determinants of creativity mentioned above. By extension, I would argue that these same factors might affect noticing, as well.

Work on social facilitation (e.g., Zajonc, 1968) suggests that the presence of others can lead to a decrement in performance at learning tasks. The effect that is heightened by the threat of evaluation, a fact that makes Amabile's (1988) observations about organization even more salient.

Summary

This section suggests that there are individual, organizational, and system-specific influences on noticing. Some are difficult for an individual to influence, for example, the structural characteristics of the system being modeled and the nature of their important undistinguished variables. Others, such as problem framing, cognitive mode, and what Amabile (1988) terms "creativity-relevant skills", suggest that we might through careful monitoring of these factors and training be able to affect noticing positively. Work on debiasing (Fischhoff, 1982), however, conveys the difficulty of altering the natural course of human reasoning.
Chapter 2

The independent variable studied in this dissertation, explication, was not identified as a "missing link" in the literature reviewed here. Rather it emerged from empirical observation and exhibits several important characteristics. First, it makes sense in light of this literature but has not been proposed as an influence on noticing. Second, it ties into ongoing debates in the psychology literature and relates to work that is ongoing in other fields as well. Finally, it is a behavioral variable that managers potentially can influence and thus offers the possibility of practically useful results. In the remainder of this chapter, I describe the formation of my central hypothesis and then develop the independent variable.

2.3. Pilot Testing and the Emergence of Explication as a Variable

My research question, or more specifically the independent variable, emerged from observations made during pilot testing. Thus it makes sense to discuss the pilot briefly as part of the background to the experiment rather than as a part of it. In this section, I discuss the pilot test and its serendipitous outcome without providing the details of the experimental design or the operationalizations of variables. The experiment is described in detail in Chapter 3. The goal here is to focus attention on the emergence of the research question during piloting.

2.3.1. Goals at the Pilot Stage

The primary goal of the experiment was to create a task and a context in which noticing could be studied. This calls for a task and context where there are opportunities for people to notice variables, where the fact and timing of noticing can be measured in a reliable fashion, and where there is variation in whether people notice or not. This last criterion implies that the task is neither so simple nor obvious that every study participant notices all there is to notice nor so difficult that no one is able to make any progress; there should be no ceiling or floor effects on the dependent variable. The task developed was a PC-based management simulator in which the participants made repeated decisions. If participants noticed two "non-obvious" variables related to advertising they could improve their ability to make predictions and thereby perform better.
Although the explication-noticing hypothesis was generated during the process of pilot testing and became the central element of this dissertation, the experiment as originally conceived was intended to test several ideas based on the literature reviewed above. The original hypotheses related framing to noticing and to the way people carried out problem-solving. In subsequent chapters, the issue of framing will reemerge as important. In the laboratory and during the piloting process, the effect of explication appears to have dominated any framing effects, creating an interesting and novel direction for the research.

The goal of the pilot was to implement the experimental design and then "tune" the experiment as necessary for it to work smoothly. Thus the pilot addressed several key questions:

- Is there reason to suspect a floor or ceiling effect on the outcome variable? Is configuring the simulation one way or the other (e.g., with the advertising data on the screen versus accessible via a button) associated with such an effect? This question really amounts to whether the experiment provides an opportunity for people to notice and an environment in which to observe variation in responses.

- Are the instructions really self-explanatory? Do people interpret them readily and can they execute the task without additional input from the experimenter?

- Do people draw any unintended interpretations from particular parts of the instructions that guide their problem-solving behavior in idiosyncratic or, for my purposes, counterproductive ways?

- Can people learn to use the simulator readily so that they are not overloaded just with the logistics of using it?

- Is the task too easy or difficult? Can people in the sample population handle the amount of information on the screen?

- Are there any problems with the way the simulator is parameterized? Should the variables to be noticed have stronger effects? Weaker?

- Should the task involve fewer decisions? More?

- How much time do people need to complete the task?

- Do participants understand the way they are scored? Is it confusing to them or do they tend to forget part way through the simulation?
2.3.2. The Birth of the Explication-Noticing Hypothesis

The experimental procedure is described in Chapter 3 and was followed closely in the pilot. The primary difference lies in the use of non-random assignment in the use of the written protocol. Initially, I was not planning to use a protocol. When the first two participants paid absolutely no attention to the advertising variables – which were displayed continuously on the screen – I decided to employ a protocol to collect process data. Participants were asked to write down their reasoning while they performed the task; the instructions asked participants to record their hypotheses – even pure hunches – and the variables they included. I wanted to understand their behavior and thinking better in an effort to improve or modify the task. I used a written protocol for the next four subjects, all of whom noticed the advertising variables. At this point, suspicious that the use of the protocol might be affecting people's performance, I ran three more people without the protocol. All three failed to notice the advertising variables. It seemed that the primary difference between the people writing the protocol and those who were not was that the protocol-writers were articulating their hypotheses; they were making them explicit. Thus the explication-noticing hypothesis was born.

During the course of the pilot work with these 9 participants, other changes were made, including the following:

- The instructions were improved based on questions that people asked.
- For some, the advertising variables were displayed continuously on the screen. Others viewed them by clicking on a button.
- Additional checks were put into the program to catch if participants typed in very high or very low prices that most likely represented typographical errors. These checks questioned high or low inputs but did not prevent participants from putting in these numbers if they confirmed their choice.
- The parameterization of the simulator was changed slightly so that halfway through all participants would experience an increase in the variance in the outcomes. Someone who had not noticed might observe the increase in overall variance (quite clear on one of the simulator's graphical outputs), become suspicious that something had changed, and search more energetically for a new variable to help explain the change. The idea was to give non-noticers a nudge.
- The original experiment used two slightly different versions of a case in orienting study participants to the simulation. These cases represented the original experimental treatment and represented an effort to manipulate their framing of the problem. People were assigned to the two versions on a random basis.

The pilot study was halted after the ninth participant. By then, several points were clear, with respect to the original goals of the pilot study:

- There was variance in noticing, which suggests that overall the task was neither too difficult nor too easy.

- By the end of the pilot, the instructions had been improved to the point where they were self-explanatory. People could work independently through the entire simulation.

- Participants did not appear to have any odd interpretations of the instructions or the case that were in any way troublesome or led to unexpected behaviors.

- People became comfortable with the logistics of using the simulator very quickly; by the second turn there was no hesitancy in entering prices and there were no questions about logistics. The biggest problem arose from the fact that Excel distinguishes between the "Return" and "Enter" keys and that only one worked properly in the simulation. I applied a low-tech fix by putting a "post-it" on the return key that said "Use the Enter key on the numeric keypad ⇒."

- Participants did not report having difficulty with the amount of information on the screen or understanding its parts once the axes of the graphs were clearly labeled.

- While there were no problems associated with the parameterization, I chose to have the parameterization change part way through the simulation, as described above.

- The task took between 30 and 60 minutes. There seemed to be no obvious need to make the game shorter. Noticing, when it occurred, happened in the first half of the game. Furthermore, there was no basis to suspect that the non-noticers would have been likely to notice had the game been just a little longer. Thus there was no obvious need to lengthen the simulation.

- All participants understood the scoring system; they all were able to describe how their score was calculated. In doing so, most noted that since it did not account for inventory carrying costs it was not a good proxy for profit. This clarification demonstrates clear understanding. When asked if they had trouble keeping this particular objective in mind as their goal, they indicated no difficulty.
Chapter 2

- Post-test interviews and behavior during the simulation did not suggest that the different versions of the case had an effect on noticing. There was no evidence that the two versions of the case were inducing different framings of the problem.

With the exception of my doubts about the effectiveness of the cases as a manipulation, the goals for the pilot test had been met. In addition, I was highly suspicious about the role that the protocols were playing. The use of the protocol was confounded with other changes made during the piloting process; nonetheless, it was perfectly correlated with noticing. After a literature review that helped me to conceptualize the writing of the protocol as the construct "explication," I decided to put the framing hypotheses on hold temporarily and pursue the experiment reported in the next chapter. In the next subsection, I define the term explication and develop the construct.

2.4. Explication

2.4.1. A Definition of Explication

Argyris and Schön (Argyris & Schön, 1978; Schön, 1983) note that the means by which the brain directs action – "theories-in-action" – are distinct from "espoused theory", the explicit symbolic representation of this knowledge that we use to explain our behavior or decisions.

Argyris and Schön's argument is consistent with many other theorists' observation that people have only limited awareness of their theories-in-action. Often, for example, work in Social Judgment Theory demonstrates that people espouse theories that are more complex than, and do not correspond to, their theories-in-action (Balke, Hammond & Meyer, 1973; see review by Slovic & Lichtenstein, 1971); typically, this work focuses on our inaccurate assessments of the weights variables assume in our judgments. Sometimes people are unaware of important variables to which they respond unconsciously or automatically. Polanyi (1958), using the term "subsidiary awareness," calls upon the idea of tacit knowledge which, through practice, has become embodied in skills and is no longer available at the conscious level. Bruner notes people's ability to act based on knowledge before they can verbalize it (Bruner et al, 1956; also Hull, 1920). Vickers (1978) argues that even though people can identify and describe deviations from norms, they have great difficulty describing them (Vickers, 1978). This work suggests that
although our theories-in-action may be well developed, we can capture them only partially in our espoused theories.

Other researchers have observed that our espoused theories can also contain elements that play no role at all in our theories in action. Of course, this is a central tenet of Argyris and Schönen's (1978; Argyris, 1994) work. They argue that our espoused theories often are essentially unrelated to our theories in action, consisting of different variables entirely. This discrepancy creates huge problems for communication and learning within organizations. The discrepancy between our knowledge and our awareness of it has also been observed in relation to attitudes. Wilson, Dunn, Kraft and Lisle (1989) note that people are often quite unaware of the bases for their attitudes and report causes that clearly were not influential at all. In discussing his work as a psychotherapist, Leston Havens mentions how "limited and inaccurate" is the process of self-reflections and wonders, "Is the mind under examination a quisling exposing its own secrets or a patriot staunch in its own defense?" (Havens, 1989: 54-55). Havens' question is reflected in an ongoing debate about the accuracy of people's descriptions of their own thought processes (Ericsson & Simon, 1980, 1994; Nisbett & Wilson, 1977; Payne, Braunstein & Carroll, 1978; Russo, Johnson, & Stephens, 1989; Smith & Miller, 1978). The care with which proponents of protocol-use specify the conditions appropriate for its use and the widespread nature of research suggesting protocols' inaccuracy make it clear that under many circumstances people are unable to represent their theories-in-action both explicitly and accurately.

Explication is the process of rendering ideas into communicable form. This research focuses specifically on the use of language, despite the fact that I can think of other symbolic representations of knowledge to which the findings might generalize. When people explicate, they convert theories-in-action – which need not be, and I argue are not, represented as language – to an explicit representation in language (see Pinker, 1994, Chapter 3, for a discussion and review on the distinction between thought and language). As individuals, people can, and often do, function quite well without explicit language-representations of their theories-in-action. Action does not require it; the presence of theory-in-action is both necessary and sufficient. Without
external requirements or incentives for explication, people might never make this jump from idea to explicit representation.

2.4.2. The Role of Explication in Cognition

Because of imperfect access to our theories-in-use, explication is not just a matter of translating already-present and clear ideas into symbolic form; it might be characterized better as an inference process, one that requires search and creativity (Wilson et al, 1989). One result of this inference process might be an increased likelihood of noticing new variables.

History offers conspicuous examples of people using private explication, typically journals, to facilitate noticing. Writers of both fiction and non-fiction have used diaries to help stimulate thought (see Rainer, 1978, for a wide-ranging discussion of the many uses of personal diaries). More germane to this discussion, scientists have long acknowledged the role that journal-keeping can play in generating and synthesizing ideas. Charles Darwin is one of the most famous examples. The use of explication, whether in the anthropologist's field journal or the biologist's lab notebook, is central to the scientific method.

Explication is also an important tool in psychotherapy (Corsini & Wedding, 1989). Through explication, patients do more than state what they already know; the process helps patients identify relevant variables, some of which they were previously unaware (Havens, 1989), to incorporate those variables into their explicit thinking, and ultimately to integrate them into their theories-in-action (see Rogers' [1961] discussion of "significant learning in psychotherapy and education").

Researchers in the field of organizational learning study the role of explication. Argyris (1994; Argyris & Schön, 1983) sees explication essentially as a matter of translating one's theories-in-action combined with honest expression of the results. Explication is viewed as important, but is not seen as a creative or inferential process. Others who study organizational learning intervene in organizations by having participants explicate as a first step toward integrating their models and resolving conflicts (Senge, 1990; Sterman, 1994; see also Lee, Courtney & O'Keefe, 1992, as a good example of
an information processing approach to organizational learning). The primary role identified for explication in this work, however, is translation. Noticing results from the interaction of people with conflicting models, not through individual-level cognitive processes evoked during explication.

Work in the area of debiasing suggests that certain types of explication may have important and useful effects on the judgment process. Fischhoff (1982) suggests that a recurring theme in efforts to debias the judgment process is "forcing respondents to express what they know explicitly rather than letting it remain 'in the head'" (Fischhoff, 1982: 427). By asking their subjects how they might explain outcomes that do not occur, Slovic and Fischhoff (1977) are able to combat hindsight bias. By asking their subjects to list reasons why their judgments might be wrong, Koriat, Lichtenstein and Fischhoff (1980) are able to reduce overconfidence. In these experiments a type of directed explication, articulating the reasons for or against the occurrence of an event, was used successfully to alter the judgment process.

Work on both sides of the debate on the validity of verbal protocols is consistent with the role I hypothesize for explication (see Ericsson & Simon, 1980, 1994; Nisbett & Wilson, 1977; Payne, Braunstein & Carroll, 1978; Russo et al, 1989, Wilson et al, 1989). While these authors disagree on the ultimate accuracy of verbal explication as a measure of people's thought processes, they acknowledge several ideas important to the current study. In the process of verbalization, people may draw information initially from short-term memory. However, if the material they are asked to produce is not currently in short-term memory, they then begin to infer what it is that they think based on long-term memory and their perceptions of their environment; through inference, people actively create new mental models to explain themselves. In the terms of the debate, this inference process creates a methodological problem. However, by framing it this way, these authors neglect the positive role that explication may play as an engine for creative processes.

The work described here suggests that explication is likely to have an effect on the thought process. Because of our poor access to our own mental models, explication involves us in an inference process. As we engage in this process, we create new, alternative explanations, form links between ideas.
that were not present before, and potentially notice new variables. Explication calls us away from scripted, routinized, and automatic thought. In so doing, it creates the opportunity for us to generate new ideas.

2.4.3. Explication in the Context of an Organization

It seems reasonable to expect that the circumstances under which explication takes place would affect its relationship with noticing. Argyris' (1994) work suggests that organizational norms driving what he calls "defensive reasoning" constrain the direction of and even halt inquiry into specific realms of organizational life (e.g., those that are threatening or embarrassing). Any explication in such a climate might also conform to the rules of social legitimacy.

Wilson et al (1989) subscribe to this notion, suggesting that when people explain themselves, they feel compelled to come up with answers that sound reasonable to others. Their inquiry acquires a constraint: it focuses on ideas that will appear legitimate to others. The need to justify one's ideas can be influential in many respects. Staw's experimental work on escalating commitment (Staw, 1976, Staw & Fox, 1977; Staw & Ross, 1978) links rigid thinking, in the form of commitment to a failed status quo, to the need to justify past decisions. In her work on creativity, Amabile (1988) finds that an organizational climate that emphasizes criticizing ideas and evaluation deters creativity on the part of its individual members.

It should be quite clear that although explication might be a private experience, perhaps used as a tool for personal exploration, learning, or creativity enhancement, many of the circumstances under which it might occur are likely to involve other people and thus be laden with social meaning. It might be a medium through which mentors teach their apprentices, a way to justify and gain approval for decisions, or an element of an organization's personnel evaluation process. Potentially, its public or private nature and the end that it serves could lead to different results either as different amounts of noticing or noticing of different types of variables.
3.5. Summary and Research Questions

Three questions guide this research. The first and most fundamental asks:

Does the process of explication affect thought in such a way that noticing new variables becomes more likely?

This section has defined noticing and explication, and proposed a causal relationship between the two. The hypothesized relationship is not inconsistent with existing literature, but neither has it been proposed in the past. The literature offers suggestions of why explication might have its effect. Primary among them is the idea that since we have poor access to our own mental models, explication involves inferring our beliefs, and therefore offers us the opportunity to introduce new ideas and challenge old ones. However, this tentative explanation drives a second research question:

What are the mechanisms through which explication has its effect on noticing?

The literature that uses explication as a construct also suggests that if an explication-noticing link exists, it is likely to be contingent on the social context within which explication takes place. Literature related to noticing, creativity, and flexible thinking also points to the importance of context. Given my concern with noticing within organizations, the dissertation's third research question takes the form:

What are the organizational characteristics likely to affect explication and moderate its relationship with noticing?

The next three chapters address these three questions in sequence. Chapter 3 presents a laboratory experiment that tests the relationship between explication and noticing. It controls as much as possible the individual- and organizational-level variables that might also have an influence. Chapter 4 presents a primarily qualitative analysis of the experimental results that identifies and demonstrates quantitatively the mechanisms through which explication works. Chapter 5 uses a field study to explore the variables present in organizational settings that can affect the process of explication and its relationship with noticing.
3. The Link Between Explication and Noticing

3.1. Overview

The previous chapter offers a theoretical perspective on the relationship between explication and noticing and describes the process that led to the generation of the research question. This chapter describes an experiment to test the explication-noticing hypothesis. It is divided into five subsections. This overview is followed by a description of the experimental design. The next subsection presents the experimental results. The chapter concludes with a discussion of the results. A qualitative discussion that examines in more detail the specific strategies and conceptual blocks that study participants experienced follows in Chapter 4.

Sloan Master's students volunteered to work with a PC-based management simulation. Their task involved making 40 repeated decisions and seeing the simulated consequences of each, thereby providing the opportunity to learn from experience. The dependent variable, noticing, was measured by whether study participants used peripheral but useful variables in their decisions. These variables related to advertising and are referred to below as "the advertising variables." Each participant was randomly assigned to one of two conditions. The treatment group received instructions intended to induce written explication; the control group did not receive these instructions.

3.2. Experimental Design

3.2.1. Task

Goals of the Task Design

The primary goal of the experiment was to create a task and a context in which noticing could be studied. This calls for a task and context where there are opportunities for people to notice variables, where the occurrence and timing of noticing can be measured in a reliable fashion, and where there is variation in whether people notice them or not. This last criterion implies that the task is neither so simple that everyone notices all the variables nor so difficult that no one is able to make any progress (no ceiling or floor effects on the dependent variable).
Additional criteria included feasibility, ease of modification, ease of use, and realism. Feasibility meant a task that could be designed and implemented on a virtually non-existent budget with limited personpower. Ease of modification seemed critical since changes would need to be made during the piloting process. In addition, I wanted to be able to use the same apparatus in the future for experiments that might differ somewhat. It was important for participants in the study to be able to execute the task with only a brief set of instructions and that all learning other than noticing would be accomplished very rapidly and thus not be a potential confounding factor. The final issue is realism. In a conversation in 1993, Josh Klayman expressed regret that the task he used in his 1988 work was so abstract. He and Ellen Langer (personal conversation, 1994) both recommended creating a task that related closely to real-world phenomenon of interest. Although there are arguments that can be made for a completely abstract task (e.g., people do not bring any mental baggage to the task), I decided on a task with some face validity as a type of decision that people would make in the working world.

The Simulation

In collaboration with a Master's thesis student (see Elmer, 1994), I designed a PC-based management simulation in which study participants make repeated decisions about pricing a single product. For each of 40 simulated weeks they are given a sales target and are asked to set a sales price for their product so as to receive orders to match the target. They are penalized equally for carrying inventory and for missing orders; they need not worry about inventory carrying costs or issues of overall productivity. Predictive accuracy is their only goal.

The repeated decision-making task follows in the tradition of Klayman (1988), Sterman (e.g., Diehl & Sterman, 1995; Paich & Sterman, 1993; Sterman, 1989a, 1989b), Wason (1960), Bruner (Bruner et al, 1956), and the social judgment theorists (e.g., Balke et al, 1973; Hammond et al, 1966, 1973; Todd & Hammond, 1965; Hagafors & Brehmer, 1983). The computer automatically records each decision made and the time at which it is entered. The noticing of a variable may be inferred by examining its correlation (through regression analysis) with the decisions made.
Using the Simulation

Figure 3.1 shows the screen provided by the simulator\textsuperscript{1}. The person playing this particular game was about to enter the price for week 15 and had clicked on the advertising button so that the advertising information is displayed.

Each simulated week, the program updates and displays the following information (clockwise from the upper right of Figure 3.1):

Demand Window

- A graph of past demand (price selected plotted against actual orders received for the product): There are three initial reference points to start with. Subsequent points appear in a different color, except for the most recent, which appears in a contrasting color so that it may be readily identified against the background of other points.

Advertising Info Window

- Information on the television advertising slots to be used this week there are four categories based on time of day (either primetime or daytime) and day of week (either weekend or weekday). Following a convention used for all screens in the simulation, labels appear in one color, data in another. Depending on configuration, this information is either accessible by clicking on the "Advertising" button or it is displayed continuously.

Performance Window

- A graph of past performance (difference between their sales target for the week and the number of orders actually received) plotted against time.

Input Data Window

- A target quantity to sell based on how much new stock came in this week and how much was left over from last week.

\textsuperscript{1} The image shown here is in black and white whereas the game uses colors. The plain text used for labels here appears blue in the game. Data, shown here using bold characters, appears black in the game. The different types of dots on the demand graph appear as different colors, per the game's instructions.
Chapter 3

Input Data Window (continued)

- A suggested price based on the target quantity and a straight line regression through past demand
- Data on the prior week including the price chosen, the target quantity, the number of orders received, and the difference between the actual orders and the target.
- Cumulative score so far in the game.
- A button to examine the advertising variables and a second button to quit the game.

Based on these data, the decision-maker enters a price. At that point, the computer calculates the number of orders received and updates the charts and data on the screen. Based on the new stock arriving this week and any left in inventory, it sets a new sales target. Finally, it computes a straight line regression through the points on the demand graph and uses it to calculate a new "suggested price" based on the sales target. The participant may then make his or her next pricing decision.

Configuration

The game can be run either with advertising accessible via a button or with the advertising variables displayed continuously on the screen. Although both configurations were used in the pilot, with three exceptions all participants in the experiment were run with advertising accessible via the button.

Despite the pilot results, it still seemed implausible that participants would ignore the advertising variables if they were on the screen. These three participants were run largely out of curiosity, to see if the behavior seen in the pilot would recur. In all three cases it did. Including the participants in the pilot group, a total of five study participants, all in the control condition, did not notice advertising when it was displayed on the screen. These three participants are maintained in the sample because, if anything, having advertising on the screen should increase its salience. Had these people noticed, it might have been attributable to the configuration. However, they did not notice. Since configuration cannot help explain their behavior, they were kept in the sample and are not differentiated from other non-noticers. As noted in the results section, a sensitivity analysis reveals that the results
reported in this chapter are not altered by dropping the three. Configuration is not used as a variable in the analyses reported below.

Parameterization

The quantity of new stock arrivals each week is a random sequence from a flat distribution ranging from 500 to 5000 and was calculated in advance. Everybody experiences the same sequence.

Demand is calculated as a linear function of price, advertising time (daytime vs. primetime), advertising day (weekend vs. weekday), and an error term, also a random sequence calculated in advance. Everybody experiences the same random errors, a uniform distribution symmetrical about zero and independent of (i.e., uncorrelated with) new stock arrivals and the two advertising variables. The vast majority of the variance (87%) in received orders is the result of the linear relationship between price and quantity. Of the non-price-related factors, advertising time accounts for 9% of the total variance and advertising day for 2.5%. The error term accounts for only 0.17% of total variance. The effects of the advertising variables were purposefully made large in comparison to the error term so they would be discernible despite the noise. Given this parameterization, it is reasonable to expect that data supporting a correct hypothesis regarding the role of either advertising variable would be readily available.

Implementation

The game was implemented using macros in Excel version 4.0 for the MacIntosh\(^2\). The simulation was implemented and piloted jointly with a Master's student (see Elmer, 1994), a skilled Excel macro programmer who worked with me as part of his Master's thesis project.

3.2.2. Experimental Treatment: Written Protocols

Written protocols, originally used as part of the piloting process, were the treatment in this experiment. Participants assigned to the treatment condition were given ten pages on which to maintain a written protocol.

---

\(^2\) Excel is a trademark of the Microsoft Corporation. MacIntosh is a trademark of the Apple Corporation.
Instructions (included in Appendix C along with a sample page from the protocol) asked them to: "Please use these pages as a diary for your thoughts as you progress." They were asked to write their ideas and hunches about how demand works in the game, what variables might make a difference, the strategies they use, exclamations, and "anything else you think we should know about what's going on." The instructions also say: "We really want to know what you're thinking about... On the other hand, don't dream up phony stuff just to tell us." Each page in the protocol corresponded to a four-week period in the game and included two scale items about confidence; the main purpose of the scale items was to give participants something easy and structured to do to ensure that they did not forget about the protocol or fail to move through it over time. People in the control group performed the decision-making task without maintaining the protocol. They did have access to writing materials but no cue was given to use it other than the request to fill in scale items in the instruction packet (which at least got them to pick up and use the pen).

Initially, participants were assigned the treatment randomly, based on a pre-calculated random sequence. Within the control group some people recorded data spontaneously while others did not. Because of this, I increased the probability of assignment to the control group so that I might have enough observations to compare three distinct groups: control subjects who did record, those who did not, and subjects in the treatment group. Although the probability of assignment to treatment groups was not 50-50, the assignment process was random.

3.2.3. Procedure

As part of the process of recruiting, volunteers were informed that the entire process would take 90 minutes and were scheduled into 90 minute time slots. This duration was chosen based on results from the pilot study.

Participants were escorted individually to an office where they were seated at a desk with a PC running the simulator. Participants were told that they were going to read a case and a set of instructions, play a management simulation game in which they make 40 decisions, and then be interviewed.
and debriefed. After they read and signed an informed consent form\(^3\), they were given the instructions, a copy of the case, and, for subjects in the treatment group, the pages for the written protocols.

Participants were told that the instructions should be self-explanatory but to feel free to ask any questions. When subjects occasionally asked questions that went beyond logistics, I generally responded, "Do what makes sense to you given the instructions and the case." In general, subjects performed the task unassisted. Finally, they were told that the schedule should provide them with enough time that they need not hurry.

The written instructions asked participants to fill in several scale items measuring their familiarity with PCs and with basic economics and then directed them to read a case that puts the game in context. The case informed them that they had been hired by a hypothetical company and that an important part of their role was setting prices on a weekly basis for a particular good that they sell, in this case plastic fashion watches. It described the company and its business, the participant's role in it, the importance of their job, and how they would be evaluated. The simulation was presented as a way of getting used to the market's behavior prior to making real pricing decisions. The case is included as Appendix A.

The case portrayed the advertising variables in passing as peripheral to the task. The narrator in the case (their hypothetical boss) told participants that: "The spreadsheet also helps me out in calculating advertising costs; but I don't want you to worry about costs yet. You should focus on accuracy." The instructions, however, were explicit; they specified that the data that comprise "advertising cost information" were two binary variables describing the advertising slots assigned for the week, one for advertising time (primetime vs. daytime) and one for day of week (weekend vs. weekday).

After reading the case, participants read a set of instructions describing the simulation, the data it presents them, how to use it, and their objective (i.e.,

---

\(^3\) All volunteers agreed to participate at this point. All completed the procedure and in interviews none claimed to have felt any desire to withdraw at any point in the procedure. The informed consent form is included as Appendix D.
how scoring works). These instructions were designed to be self-explanatory in order to minimize contact between the experimenter and participant prior to completion of the simulation exercise. The instructions are included as Appendix B.

After reading the instructions, they worked with the simulation. Once they had completed the 40 time periods of simulation, they were interviewed and debriefed.

**Means of Data Collection**

Data were collected in three ways:

- Questionnaire items were used to measure background variables such as familiarity with personal computers. The questions were included in the instructions and participants marked their answer on scales printed immediately below the questions.

- For each decision made in the simulation, the computer automatically tracked the price chosen, the timing of the decision, and whether the person looked at the advertising data.

- Finally, semi-structured interviews were used after the simulation exercise to help measure the variables people used in the simulation, whether they had noticed the advertising variables, and their motivation. I asked them what information they had recorded and why, whether they had felt time pressure, and about their backgrounds. The interviews were recorded and transcribed.

**3.2.4. Sample**

The sample consisted of 42 second-year Sloan Master's students who volunteered to participate in the study. They were recruited primarily by making announcements in classes in which I described the study briefly, asked for volunteers, and passed around a sign-up sheet.

Sloan School students were chosen for several reasons. First, they represent tomorrow's mid-level managers, a population whose role in the innovation process is of interest to the author. Second, they were readily available; recruiting them was a straightforward process. Finally, the students
already have the skills necessary to understand the simulation readily and learn the logistics of the user interface in just a few minutes.

The fact that the sample consists of volunteers raises the possibility that they might differ in some important way from the rest of the student population. I address the generalizability of these findings later, in discussing the results. Since random assignment was used in the experiment, the nature of the sample raises no issues regarding internal validity.

I discovered after all data were collected that three participants were undergraduates, not Master's students. Since I felt that these participants were quite different from the rest, I removed them from the sample. However, I checked all analyses to see if any were sensitive to including this group in the sample. In no case did their inclusion change the result of an analysis.

Eight members of the sample were enrolled in a class in system dynamics. I was concerned that this group might behave differently from the rest of the sample because of their involvement in a course focusing on systems and their analysis. I checked the random assignment sequence once these participants were scheduled. Both experimental conditions were proportionally represented; in effect, the sample was stratified and I did not need to alter the pre-calculated random assignment sequence. One participant (a non-noticer) commented on how his experience with system dynamics made him curious and motivated during the simulation task. Participants offered no other references to it and analysis of a dummy variable for system dynamics does not improve the fit of the models in the analyses presented below.

3.2.5. Variables and Measures

Noticing

Several approaches were used to measure noticing. They are described below:
When people looked at the advertising data

For almost all participants\textsuperscript{4}, the simulation was configured so that the information to be noticed could be seen only by clicking on a button. In this mode, the program recorded in which time periods of the simulation they clicked on the button. Thus I knew, for example, when each participant first looked at advertising and how consistently they looked at it from that point on. The computer also recorded the time for each turn but could not capture time spent looking at the advertising.

Post-test interviews

The interview asked about the variables used for making decisions after playing they game. Initially, I asked open ended questions about the variables participants considered, the strategies they used, and how their thinking evolved over time. After exhausting this line of inquiry, I would ask more closed-ended questions about variables that they had not mentioned or other specific, unresolved issues.

Throughout this discussion I asked participants to distinguish between ideas that had crossed their mind during the simulation, ideas that they had actually acted on during the simulation, and ideas that were occurring to them only in the interview.

Inference based on decisions: Regression-based approach

The computer recorded participants' decisions and the value of the variables in the simulated environment. These data were analyzed to infer use of the advertising information. There are several approaches that one might have used to make such inferences. I used two, one based on regression analysis and a second chart-based approach.

The regression based approach follows in the tradition of Klayman (1988), Sterman (1989a, 1989b; Diehl & Sterman, 1995; Paich & Sterman, 1993), and the Social Judgment Theorists (for example, Hammond, Summers & Deane, 1973; Todd & Hammond, 1965). In this approach, the values of decisions are regressed on the values of the cues that make up the expected set of

\textsuperscript{4} See the subsection on configuration in Section 3.3.2 for a discussion of the implications of
predictors. In much of this work, the cues (i.e., the independent variables to be used in the regressions) are developed in advance. In this study, the design of the task dictated some (the suggested price\textsuperscript{5} and the two dummy variables representing the two dimensions advertising). Through the interviews, I learned of other variables that participants thought might be important and included them, as well (prior week's price, prior week's performance). A number of participants hypothesized non-linearities in the price-quantity relationship; however, their hypotheses were too idiosyncratic to justify including the required set of predictors in all regressions (target quantity\textsuperscript{2}, target quantity\textsuperscript{3}), particularly given the relatively small number of observations.

The regression analyses were based on the participants' last 25 decisions made in each 40-turn game. Using more observations captured too much confusion from the beginning of each game. Using fewer offered no advantage since almost all noticers were beginning to use advertising by the halfway point in the game.

If the coefficient for a given variable was statistically significant (p<0.05), this measure would suggest that the variable had been noticed (and also used consistently). The p-values for all participants are reported in Table 3.2 and discussed in the results section.

Inference based on decisions: Chart-based approach

The regression approach has a weakness, however. Participants who test several hypotheses involving advertising, and thus do not use a single approach consistently, are likely to be identified as non-noticers in a regression-based analysis. The chart-based approach allowed me to verify by eye (admittedly, a less-than-rigorous test) claims made by participants in the interview about strategies that they employed for brief periods of time. For example, one participant claimed to have priced high relative to the suggested price, then priced low, then stuck with the suggest price, and then priced high or low based on his success in the prior week. A strategy such as this or any

---

\textsuperscript{5} As mentioned earlier, the suggested price was calculated based on the target sales quantity and a straight-line, least-squares regression through the past demand data.
strategy adopted late in the game would be missed by the regression-based approach but could be readily checked using a chart such as Figure 3.2.

**FIGURE 3.2**

*Deviation from Suggested Price, Coded by Advertising*

Figure 3.2 shows one participant's decisions throughout the game. The horizontal axis represents time. The vertical axis is the difference in dollars between the price quoted and the price suggested by the computer. Almost all participants used the suggested price as an anchor in their decision-making so it generally made sense to use charts of this form\(^6\). The data points for each

\(^6\) One participant (#16) claimed to have anchored off of a rough mean price of $26 rather than the suggested price. To establish if this person had in fact used the claimed strategy, I plotted the difference between price quoted and $26. This plot showed a clear pattern consistent with the strategy described in the interview. This person's claim is also consistent with the regression results. (see Table 3.2).
week are color-coded on the basis of the two advertising variables. Figure 3.2 suggests that this participant began responding to the advertising variables consistently somewhere around week 23 or 24.

Summary

These four approaches to measuring noticing – monitoring when people looked at the advertising variables, post-game interviewing, using regression analysis to test the relationships between variables and participants' decisions, and using chart-based visual analysis – together do much to compensate for the weaknesses of each individually. The first measures whether participants actually looked at the advertising information but not whether they distinguished useful variables. The interviews can provide extensive data but may fall prey to memory biases or post hoc reconstruction. The regression analyses provide convincing evidence of associations; however, given enough non-noticing participants, there will be some Type-I errors (i.e., false-positives). In addition, the regression captures noticing only if the participant realized a consistent strategy. The chart-based approach is the least rigorous and yet is the only way to verify certain of the participants' claims made in the interviews. Together, these four methods can be used to paint an accurate and very descriptive picture of whether participants noticed the advertising variables, when they did so, and what hypotheses they considered about them.

The coding scheme for noticing was developed empirically from these data and is discussed in the results section.

Recording Behavior

During the simulation, study participants recorded various types of information. People in the control group used the backs or the margins of their instructions. They were already in the habit of writing since most (all but 9 people) underlined parts of the case or took notes in the margins as they read it. People in the treatment group used the pages provided for the protocol but also often drew pictures or jotted down numbers on their instructions.
The approach to coding these behaviors and the content of the information recorded was developed empirically and is described in the results section.

Time

The computer recorded how much time people spent on each turn. It did not record time spent reading or thinking about the case. The analyses below generally use the total time spent playing the game (in minutes) as a variable. However, some use the time spent on the first five turns of the game to capture more accurately participants' behavior early in the game.

Strategies, Hypotheses, and Variables Used

In the interviews, I asked participants to describe their decision-making strategies. Whenever possible, I verified these descriptions by examining the participant's decisions using tools such as the chart shown in Figure 3.2. In all cases I also regressed the decisions on specific variables. A good example of this was when a person claimed to have consistently based his or her decisions on the previous week's performance; such a claim was easily checked via regression. In the best cases, this verification provided solid evidence of their descriptions. For example, some people felt that there was a feedback loop from prior performance to current demand and claimed to have adjusted their decisions accordingly. After regressing these people's pricing decisions in the last 20 time periods of the game on the target quantity, the two advertising variables, and prior performance, prior performance was always a statistically significant predictor (p<0.05). This was not the case for those who had not claimed to have used this variable in their decision-making. Some participants claimed to have held prices stable for period of time, testing the hypothesis that their price changes were affecting demand. Others claimed to have tried using "99s" (i.e., prices such as $25.99 rather than $30.00) or pricing high for a bit, then low. These claims were easily verified.

Many claims could not be verified in such a concrete manner. Typical among these were people's claims that they priced higher or lower at some point on the curve. Often there was just too much noise in their decisions to verify these claims and the problem would be exacerbated if they used the strategy for a relatively short period of time and then dropped it.
I discuss these data extensively in Chapter 4, which presents a qualitative analysis of participants' behavior in the experiment.

Motivation

In the interview, I asked people directly about their level of motivation, whether it changed during the game, and about their level of frustration. I also asked participants if there were any points in the game when they gave up or just guessed without using any strategy.

There are other variables that are likely to be related to motivation and thus might serve as proxies for it. Among these are the time participants took executing the task and the amount of information they recorded.

Time Pressure

In the interview, I asked participants if they experienced any time pressure during the simulation and whether at any point they had felt they should speed up. I regard anyone who failed to say no as someone who may have felt pressured.

Familiarity with Personal Computers

It was important for participants to be comfortable with the logistics of using the simulator. I assumed that people who used PCs frequently would be. The following question was included with the instructions: How much do you use PCs (spreadsheets or word processors)? It was answered by circling a number on a 7-point scale ranging from 1 (never) to 2 (less than once per month) to 7 (daily).

Familiarity with Economics

It was also important for participants to be familiar with the notion of a demand curve. The following question included with the instructions: How familiar are you with the basic idea of a demand curve as taught in introductory economics courses? It was answered by circling a number on a 7-point scale ranging from 1 (not at all familiar) to 7 (very familiar). It should be noted, however, that all Sloan Master's students must have taken intermediate undergraduate micro-economics to be accepted and take a required course in micro-economics in their first term. It would have been surprising if they had not reported high levels of familiarity.
Background

In the interview, I asked each participant to describe briefly his or her educational background prior to Sloan, any jobs held prior to Sloan, and any particular areas of specialization at Sloan. I also asked if people were enrolled in a system dynamics course; as I mentioned when discussing the sample, I had some concerns that this group might prove to be different. People were classified into one of three groups: 1) people with backgrounds in business, economics or finance, 2) people with backgrounds in science or engineering, and 3) people with other backgrounds. My goal in classifying people this way was to be able to separate out people with particular analytic skills and those with higher levels of familiarity with the type of problem posed by the experimental task.

Subsample

The sample was run in two separate groups at two different points in time during the semester, one at the end of a semester and the other at the beginning of the next. I maintained a distinction between these two groups because I wondered if the group run at the end of the semester would be under more stress and thus perform more poorly at the task.

3.3. Results

3.3.1. Emergent Categories of Noticing and Recording Behavior

Categories of Noticing

Four categories of outcome with respect to noticing emerged from the participant's behavior (as recorded by the computer) and the content of the interviews. Table 3.1 reports their frequencies. Some study people never looked at the advertising variables. A second group looked at the advertising variables but never developed any hypotheses about them as potentially relevant to their decisions. A third looked at the variables, developed at least one hypothesis about their role, but did not use them in any consistent decision-making strategy. The final group looked at the advertising variables and ultimately used them as part of a consistent decision-making strategy.
TABLE 3.1
Levels of Noticing and their Frequencies of Occurrence

<table>
<thead>
<tr>
<th>Looked at advertising?</th>
<th>Developed hypotheses about advertising?</th>
<th>Using adv. consistently at end of game?</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. No</td>
<td>No</td>
<td>-</td>
<td>20 (48%)</td>
</tr>
<tr>
<td>2. Yes</td>
<td>No</td>
<td>-</td>
<td>6 (14%)</td>
</tr>
<tr>
<td>3. Yes</td>
<td>Yes</td>
<td>No</td>
<td>3 (7%)</td>
</tr>
<tr>
<td>4. Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>13 (31%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>42 (100%)</strong></td>
</tr>
</tbody>
</table>

Participants in the second group all looked at advertising 1 or 2 times, those in the third looked 6-15 times, and those in the fourth looked at least 16 times. By week 30 of the simulation, and generally earlier, all participants in the fourth group had started using a strategy that exhibited a consistent directional response to the advertising variables. Among participants this group who used only one of the two advertising variables, none expressed a belief that the second was irrelevant, although for many its role was an open issue at the end of the game. Thus, noticing one advertising variable but not the other did not emerge as a coding issue.

The regression results are shown in Table 3.2, which reports coefficient p-values for regressions that use suggested price, the price quote in the prior week, the discrepancy between the target and actual orders, and the two advertising variables as predictors of the participant's pricing decisions. A statistically significant coefficient for a variable implies that the participant noticed the variable (and, in fact, was using it fairly consistently).

There was a high degree of consistency across the four measures of noticing, as reported in Table 3.3. Two participants who never looked at advertising had a statistically significant coefficient for the weekend/weekday advertising variable (one at the 95% confidence level, the other at 90%). There were three participants who looked at advertising, tested several
<table>
<thead>
<tr>
<th>Participant</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suggested Price</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Prev. week’s quote</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Prev. week’s discrepancy</td>
<td>0.09</td>
<td>0.01</td>
<td>0.06</td>
<td>0.00</td>
<td>0.08</td>
<td>0.01</td>
<td>0.08</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Primetime/daytime</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Weekend/weekday (Constant)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Lowest tolerance statistic&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.71</td>
<td>0.66</td>
<td>0.75</td>
<td>0.66</td>
<td>0.75</td>
<td>0.66</td>
<td>0.75</td>
<td>0.66</td>
<td>0.75</td>
<td>0.66</td>
<td>0.75</td>
<td>0.66</td>
<td>0.75</td>
<td>0.66</td>
</tr>
<tr>
<td>Durbin-Watson statistic</td>
<td>0.70</td>
<td>1.62</td>
<td>1.47</td>
<td>1.76</td>
<td>1.69</td>
<td>1.72</td>
<td>1.47</td>
<td>2.22</td>
<td>1.56</td>
<td>2.22</td>
<td>1.90</td>
<td>2.20</td>
<td>2.09</td>
<td>2.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suggested Price</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Prev. week’s quote</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Prev. week’s discrepancy</td>
<td>0.07</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Primetime/daytime</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Weekend/weekday (Constant)</td>
<td>0.01</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Lowest tolerance statistic&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.80</td>
<td>0.01</td>
<td>0.07</td>
<td>0.06</td>
<td>0.01</td>
<td>0.07</td>
<td>0.06</td>
<td>0.01</td>
<td>0.07</td>
<td>0.06</td>
<td>0.01</td>
<td>0.07</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>Durbin-Watson statistic</td>
<td>1.20</td>
<td>0.93</td>
<td>1.76</td>
<td>2.01</td>
<td>2.02</td>
<td>1.78</td>
<td>1.68</td>
<td>1.43</td>
<td>1.88</td>
<td>2.05</td>
<td>1.67</td>
<td>2.57</td>
<td>1.50</td>
<td>1.79</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant</th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
<th>42</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suggested Price</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Prev. week’s quote</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Prev. week’s discrepancy</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Primetime/daytime</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Weekend/weekday (Constant)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Lowest tolerance statistic&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Durbin-Watson statistic</td>
<td>1.81</td>
<td>2.34</td>
<td>1.63</td>
<td>2.49</td>
<td>1.48</td>
<td>1.59</td>
<td>1.94</td>
<td>1.24</td>
<td>1.45</td>
<td>2.02</td>
<td>1.78</td>
<td>1.59</td>
<td>N/A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.44</td>
</tr>
</tbody>
</table>

---

<a>Table 3.2</a>

Coefﬁcient P-values<sup>a</sup> in Decision-Predictor Regression Analysis

---

<sup>a</sup> To aid reading the table, coefﬁcient p-values greater than 0.10 are not shown.

<sup>b</sup> The lowest coefﬁcient tolerance statistic is reported. Only tolerance statistics below 0.87 are shown.

<sup>c</sup> The Durbin-Watson statistic for the participant who used suggested price every turn could not be calculated.
hypotheses, but never developed consistent strategies. Predictably, the regression results for these three participants yield no statistically significant coefficients for either advertising variable. For all other participants, noticer and non-noticer alike, the measures were consistent.

**TABLE 3.3**
Frequency of Agreement Across Measures of Noticing

<table>
<thead>
<tr>
<th>Consistency/inconsistency across measures of noticing</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>All four measures of noticing are consistent</td>
<td>37</td>
</tr>
<tr>
<td>Regression results inconsistent with others:</td>
<td></td>
</tr>
<tr>
<td>• Type-I error in regression results</td>
<td>2&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>• Type-II error in regression results</td>
<td>3</td>
</tr>
<tr>
<td>Other inconsistencies in measures of noticing</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>42</strong></td>
</tr>
</tbody>
</table>

<sup>a</sup> Participant #6 at the 95% confidence level, participant #37 at 90%.

In the quantitative analyses that follow, noticing is treated as a dichotomous variable. Except where explicitly stated, the variable used classifies people in categories 3 and 4 of Table 3.1 as "noticers"; the others are considered "non-noticers". I used this coding because I am most interested in the process by which people become able to articulate hypotheses; each noticer in this scheme articulated some hypothesis about the role of advertising. In some cases, their models were incorrect. However, the errors in their models did not mean that the advertising variables were any less "noticed". Since a coding scheme that treats only the last group as noticers might be reasonable as well (although not as well suited to my goals), all analyses were checked to see if their results were consistent for both codings. In no case did use of the alternate coding change the result of an analysis.

In the fourth chapter, I distinguish between participants in the third and fourth groups and explore the reasons why the three participants in the third group behaved as they did.

**Types of Recording**

The participants varied both in the nature of the information they recorded and in the amount they wrote down. Examination of what they wrote down reveals three distinct types of information:
• Theories about what was causing the simulation results and thus the basis on which they should make decisions (i.e., explication of their hypotheses about causality in the simulated environment). This information came in the form of statements such as:
  o "Must adjust for when ads are shown, but by less than the suggested 30 cents / 100 units."
  o "Demand is more stochastic than I thought."
  o "I'm trying to pay more attention to the ad info."
  o "Stuck to recommendations [suggested price]. Seeing regular oscillations around 0."
  o "I think the market is getting saturated because in week 37 I just priced well below SRP and didn't sell the stock."
  o "Primetime is very important."
  o "Above $35, demand becomes elastic once again."
  o "I don't know if the market has a memory - I think it is BAD if the price jumps all over from week to week, so I'll try to keep it somewhat level..." Then, 20 weeks later: "I don't care whether the market has a memory, i.e., I don't think it does."

• Data on events that took place (i.e., recording the values of variables that they had identified as potentially relevant). Sometimes people kept tables. These tables generally came in one of two forms. The first were simple columns of data with each row corresponding to a simulated week. A few participants organized the data into a 2x2 table with each axis representing one of the advertising variables. Sometimes, however, data would be recorded in the form of statements:
  o "Rounded up to the nearest dollar. Got screwed on sales - undersold by 609!"
  o "Prime weekday, + $1.30. Still not enough!.. Day weekend, high Q, -1.25. Still not enough!"

One might argue that these statements sometimes contain more information than pure data. I discuss this issue in the results section.

• Information relating to their emotional state, including exclamations of frustration, surprise, etc.
  o "I continue to be optimistic regarding price setting in the long run."
  o "Getting irritated. Doing worse."
  o "I gotta get out of this business... I suck at this."
  o "YES!! Getting feel for ads."
Chapter 3

The amount that people wrote fell into three categories, as well. People either:

- Recorded no information
- Recorded information on a few scattered weeks but exhibited no systematic recording of two or more variables
- Recorded information systematically for an extended period in the game

There was a clear boundary between people in the second group, who without exception took notes on 6 or fewer occasions, and those in the latter group who recorded information on at least 11 consecutive periods. Of the 5 participants who recorded data on 6 or fewer occasions, 3 had breaks in their recording (they did not record a consecutive block of time periods) and the other two recorded only a single variable. Of the 16 members of the other group, all recorded data consistently and did it for at least 2 variables (11 for 3 or more).

In the analyses that follow, the amount of information recorded is treated as a dichotomous variable. Independent of type of information, only people who recorded information systematically are treated as "recorders." Those who recorded nothing or took only a few notes are grouped together and categorized as "non-recorders." I checked the analyses to see if any were sensitive to the shift to a coding scheme that groups together the occasional note-takers and those who record systematically. In no case did use of the alternate coding change the result of an analysis.

There were two participants whose mode of recording data clearly reflected hypotheses that they had but who did not record anything else that might be interpreted as explication. The first drew a graph with price and quantity axes, put a straight line through it, and then began recording at the appropriate spot on the graph the amount higher or lower the "true" demand was relative to the straight line. Clearly this person hypothesized that the price-quantity relationship was nonlinear. The second organized the data she recorded into a 2x2 table whose axes corresponded to the advertising variables. Clearly this mode of recording reflected a hypothesis about advertising. I do not equate these decisions about recording data with explication, however. Nor do I treat as explication statements such as "Rounded up to the nearest dollar" which
might also suggest a hypothesis. They, like the modes of recording data presented above are the product of a hypothesis rather than a conscious effort to state the hypothesis explicitly, for example, "Maybe customers like 99's" [i.e., prices of the form $23.99, $28.99, etc.]. The statements and modes of recording were behavior, driven by unexplicated hypotheses. Therefore I treat them as behavior as data-recording, not explication.

The resulting coding scheme classified study participants on the basis of whether they recorded data (data-recording) and whether they recorded their hypotheses (explication). Only one person recorded emotional state systematically.

**Recording as a Response to Noticing**

Up to this point, I have discussed recording behavior as a consequence of treatment category. An examination of the timing of recording, particularly in relation to the point at which people began looking at the advertising variables, reveals interesting distinctions.

Everyone who actually explicated began to do so at the start of the simulation. However, of the people who recorded data, several began in the same time period in which they first looked at advertising or very soon thereafter. In these cases, data-recording could have been (and likely was) a response to noticing the advertising variables. However, data-recording may have played an important role in the second component of the noticing process, linking the variable to the model by moving it into the candidate set from the set of known variables. A coding scheme that categorizes these participants as data-recorders acknowledges this potential causal role of data-recording in the noticing process.

Operationalizing data-recording in this fashion also introduces a conservatism into the analysis that is desirable given the study's primary hypothesis. It increases the explanatory power of data-recording in predicting noticing because it includes all situations where it potentially is a cause and some where the two are linked but causality is reversed. Using this coding therefore makes it more likely that data-recording, rather than explication, is
likely to be identified as a predictor of noticing\textsuperscript{7}. Since my hypothesis is that explication, and not data-recording, mediates the relationship between treatment and noticing, an operationalization of data-recording that strengthens its role (or that, at least if biased, is biased in this direction) makes my analysis conservative, that is, less likely to disconfirm a role for data-recording, and therefore less likely to attribute one to explication instead.

### 3.3.2. Descriptive Data

Table 3.4, below, indicates the associations between the primary variables used in the analyses. Table 3.5 displays frequencies for the categorical variables, and means and standard deviations for those that are continuous.

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Rec. data</th>
<th>Explicated</th>
<th>Time</th>
<th>Notice</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recorded data</td>
<td>.43**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explicated</td>
<td>.89***</td>
<td>.42**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>.70***</td>
<td>.69***</td>
<td>.63***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noticed</td>
<td>.43**</td>
<td>.29~</td>
<td>.54***</td>
<td>.51**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>.51**</td>
<td>.44**</td>
<td>.59***</td>
<td>.47**</td>
<td>.59***</td>
<td></td>
</tr>
<tr>
<td>Subsample</td>
<td>.16</td>
<td>.08</td>
<td>.08</td>
<td>-.06</td>
<td>-.21</td>
<td>-.15</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Pearson correlation coefficients are used to express relationships between interval variables, Phi between binary variables, and Spearman between binary and interval variables.

### Background and Gender

The distributions of the study participants' genders and professional/educational backgrounds do not differ from those of the population of Sloan Master's students in general (for gender, $\chi$-square=2.17, df=1, n.s.; for background, $\chi$-square=1.85, df=2, n.s.).

There is no \textit{a priori} reason to hypothesize that gender would influence the results of this experiment and, in fact, it did not; it is included here only to

\footnotesize
\textsuperscript{7} In fact, this is the case. The relationship between data-recording and noticing is less strong if the participants in question are not considered data-recorders.
indicate that the sample is representative of the population of Sloan Masters students\textsuperscript{8}.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background</td>
<td>Econ/Fin/Bus</td>
<td>10 (24%)</td>
</tr>
<tr>
<td></td>
<td>Tech/Science</td>
<td>22 (52%)</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>10 (24%)</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>14 (33%)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>28 (67%)</td>
</tr>
<tr>
<td>Treatment Category</td>
<td>Treatment</td>
<td>13 (31%)</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>29 (69%)</td>
</tr>
<tr>
<td>Subsample</td>
<td>Group 1 (End of Term 1)</td>
<td>24 (57%)</td>
</tr>
<tr>
<td></td>
<td>Group 2 (Start of Term 2)</td>
<td>18 (43%)</td>
</tr>
<tr>
<td>Data-recording</td>
<td>Recorded data</td>
<td>16 (38%)</td>
</tr>
<tr>
<td></td>
<td>Did not record</td>
<td>26 (62%)</td>
</tr>
<tr>
<td>Explication</td>
<td>Explicated</td>
<td>11 (26%)</td>
</tr>
<tr>
<td></td>
<td>Did not expl.</td>
<td>31 (74%)</td>
</tr>
<tr>
<td>Noticing</td>
<td>Noticed</td>
<td>16 (38%)</td>
</tr>
<tr>
<td></td>
<td>Did not notice</td>
<td>26 (62%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to Complete (min.)</td>
<td>35.1</td>
<td>17.9</td>
</tr>
<tr>
<td>Score</td>
<td>22317</td>
<td>4384</td>
</tr>
<tr>
<td>Familiarity w/ Econ.\textsuperscript{a}</td>
<td>6.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Familiarity w/ PCs\textsuperscript{a}</td>
<td>6.1</td>
<td>1.1</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Scores are on a 7-point scale, with 7 corresponding to high familiarity

Background was included to indicate the representativeness of the sample, as well. But one might also reasonably hypothesize that it influences the approach to problem-solving that students adopt and their initial framing of the task (e.g., their propensity to include advertising-related variables in their

\textsuperscript{8} There is no association between gender and treatment condition (\(\chi^2\)-square=0.89, df=1, n.s.), as one would expect with random assignment. Gender has no association with the outcome variable (\(\chi^2\)-square=0.81, df=1, n.s.).
candidate set). There is no association between background and treatment condition ($\chi$-square=0.62, df=3, n.s.), as one would expect with random assignment; nor does background emerge as a useful predictor in the analyses below.

Familiarity with PCs, Familiarity with Economics, and Comprehension of the Task

As predicted and as shown in Table 3.5, the study participants reported very high levels of familiarity with personal computers and with the concept of a demand curve as taught in introductory economics.

The participants had very few questions about the logistics of working with the PC; the only question that more than two people asked was, "Can I retype a price if I change my mind?" They also had few questions about the meaning of the data that appeared on the screen. One of the first participants in the experiment asked if the demand graph indicated the quantity of orders received or the quantity of product shipped. The axis was relabeled from "Quantity (units)" to "Quantity Ordered (units)" to avoid any confusion.

In the interviews, all participants indicated that they understood the meaning of the demand graph, of the performance graph, and of the data presented on the screen including the suggested price. All could describe how their score was calculated and thereby demonstrated an understanding of their goal in working with the simulator.

Subsample

The participants were run in two subsamples, Group 1 at the end of a semester and Group 2 at the beginning of the next. It is reasonable to include subsample as a control variable based on three concerns. First, the end of a semester might be significantly more stressful for students than the beginning and thus the participants in the two subsamples might be under different

---

9 The answer was, "Yes, use the delete key to erase the number you have entered and then retype it."

10 There were participants who clicked on the button to look at the advertising variables but did not distinguish any information in the window that popped up.
levels of stress outside the laboratory. Second, the degree of stress that study participants experience outside the laboratory might affect their performance in the experiment. Finally, the workload might affect the type of student who volunteers.

Within the Group 1, fewer participants (29%) noticed the advertising variables than in Group 2 (50%); however the difference is not statistically significant ($\chi^2_{(1)}=1.89$, n.s.). The analysis of subsample as a control variable is distributed throughout the remainder of this chapter. To summarize briefly, it does have a main effect, but does not interact with other variables. Thus the results are consistent across the two subsamples.

### 3.3.3. Noticing

#### The Relationship between Treatment and Noticing

Table 3.6 indicates the number of participants noticing the advertising variables within each of the two treatment categories. Of people in the treatment group, 69% (9 of 13) noticed the advertising variables compared to 24% (7 of 29) of those in the control group. The difference in rate of noticing between these two groups is statistically significant ($\chi^2_{(1)}=7.74$, $p<0.01$).

<table>
<thead>
<tr>
<th>Treatment Category</th>
<th>Noticed adv.</th>
<th>Did Not Notice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment: Asked to maintain diary</td>
<td>9 (69%)</td>
<td>4 (31%)</td>
</tr>
<tr>
<td>Control: Not asked</td>
<td>7 (24%)</td>
<td>22 (76%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>16 (38%)</strong></td>
<td><strong>26 (62%)</strong></td>
</tr>
</tbody>
</table>

People in the treatment group were almost 3 times as likely\(^\text{11}\) to notice the advertising variables than those in the control group.

#### Noticing and Performance

Noticing was associated with an increment in performance ($F_{(1,40)}=20.5$, $p<0.001$). The data are shown in Table 3.7. Noticing was also associated with an increase in time spent on the task ($F_{(1,41)}=13.7$, $p<0.001$).

---

\(^{11}\) For those accustomed to odds ratios, the odds of noticing in the treatment group is 7.1 times higher than the odds in the control group.
TABLE 3.7
Mean and Standard Deviation of Score and Time by Noticing

<table>
<thead>
<tr>
<th>Noticed advertising</th>
<th>Mean Score&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Mean Time&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25664 (4811)</td>
<td>46.4 (16.4)</td>
</tr>
<tr>
<td>Did not notice</td>
<td>20386 (2683)</td>
<td>28.0 (15.1)</td>
</tr>
</tbody>
</table>

<sup>a</sup> One observation was removed from the sample for the purpose of calculating the mean and standard deviation of score and in analyses using score as a variable. This person had a unique strategy in the game and scored almost 11 standard deviations worse than the rest of the study participants. This person is included in all other analyses and is discussed separately.

<sup>b</sup> Time to complete simulation in minutes

I discuss the tradeoff between these two performance dimensions in subsequent discussions. For the time being, it is important to remember that participants were told that they would be evaluated solely on the basis of score; they were not put under any time pressure. Strictly speaking then, noticing was associated with higher performance.

3.3.4. The Role of Explication

The Treatment and Explication

Table 3.8 reports the type(s) of information that people in each of the two treatment groups recorded. In the treatment group, all participants recorded some information systematically. The vast majority (11 of 13) explicated (i.e., recorded their hypotheses) as requested. Of these, eight recorded data, as well. A single person recorded data, alone, and another recorded only information about emotional state.

TABLE 3.8
Recording Behavior, Summarized by Treatment Group

<table>
<thead>
<tr>
<th>Recording Behavior</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Asked to</td>
</tr>
<tr>
<td></td>
<td>maintain diary</td>
</tr>
<tr>
<td>Explicated theories</td>
<td>3</td>
</tr>
<tr>
<td>Explicated and recorded data</td>
<td>8</td>
</tr>
<tr>
<td>Recorded data</td>
<td>1</td>
</tr>
<tr>
<td>Recorded emotional state</td>
<td>1</td>
</tr>
<tr>
<td>Recorded no information</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>13</strong></td>
</tr>
</tbody>
</table>
Within the control group seven people spontaneously chose to use the writing materials available to them and record data. Twenty-two members of the control group did not record anything. None explicated and none recorded information about emotional state.

The request to maintain a written protocol increased the likelihood of recording information ($\chi^2_{(1)}=20.7$, p<0.001) and made a qualitative difference in the type of information recorded (Fisher's exact test p=0.00046$^{12}$). The treatment clearly did evoke recording behavior that reasonably may be labeled "explication"; this behavior did not occur spontaneously within the control group. The treatment was also associated with an increased likelihood of data-recording ($\chi^2_{(1)}=7.74$, p<0.01$^{13}$, a variable whose effect on noticing must be considered as a rival to that of explication.

**Explication and Noticing**

Table 3.9 reveals that study participants who explicated were more than three times as likely$^{14}$ to notice as people who did not explicate ($\chi^2_{(1)}=9.70^{15}$, df=1, p<0.01).

**TABLE 3.9**

Frequency of Noticing, Summarized by Explication

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Noticed adv.</th>
<th>Did Not Notice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicated</td>
<td>9 (82%)</td>
<td>2 (18%)</td>
</tr>
<tr>
<td>Did not explicate</td>
<td>7 (23%)</td>
<td>24 (77%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>16 (38%)</td>
<td>26 (62%)</td>
</tr>
</tbody>
</table>

If explication were the mechanism through which the treatment affects noticing, then: 1) it should be a better predictor of noticing than is the

---

$^{12}$ 11 of 13 (85%) of the participants in the treatment group who recorded information explicated; none of the members of the control group who recorded information explicated (Fisher's Exact test p=0.00046).

$^{13}$ 9 of 13 (69%) of the participants in the treatment group recorded data compared to 7 of 29 (24%) in the control group.

$^{14}$ The odds ratio is 15.4.

$^{15}$ Yates' correction for continuity is used.
treatment, 2) it should have an effect when treatment is controlled, and 3) treatment should lose its predictive value when explication is controlled (Baron & Kenny, 1986).

Table 3.10 allows consideration of all three of these conditions. Comparison of Models 1 and 2 reveals that the model using explication is a much better predictor of noticing than is one using the treatment. A comparison of Models 2 and 3 reveals that adding treatment as a predictor to a model that already contains explication does not improve model fit (Δlog likelihood = 0.46, df=1, n.s.). A similar comparison between Models 1 and 3 reveals that adding explication to a model with treatment does (Δlog likelihood = 4.49, df=1, p<0.05). Explication predicts noticing, even when

<table>
<thead>
<tr>
<th>MODEL</th>
<th>CONSTANT</th>
<th>LOG-LIKELIHOOD</th>
<th>SUBSAMPLE (CONTROL)</th>
<th>TREATMENT CATEGORY</th>
<th>EXPILATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL 1</td>
<td>-2.15**</td>
<td>43.57</td>
<td>1.69*</td>
<td>2.57***</td>
<td>2.25***</td>
</tr>
<tr>
<td>MODEL 2</td>
<td>-2.16**</td>
<td>39.54</td>
<td>1.65*</td>
<td>3.25***</td>
<td>9.31*</td>
</tr>
<tr>
<td>MODEL 3</td>
<td>-2.04**</td>
<td>39.08</td>
<td>1.53~</td>
<td>6.17</td>
<td>3</td>
</tr>
</tbody>
</table>

*: p<0.10  **: p<0.05  ***: p<0.01

The subsample control variable cannot be removed from Model 2 without reducing model fit (Δlog likelihood = 4.01, df=1, p<0.05); however, its interaction with explication is not a useful predictor of noticing (Δlog likelihood = 0.60, df=1, n.s.). Adding background with or without its interaction with explication to Model 2 does not improve model fit (Δlog likelihood=1.93, df=1, n.s., with the interaction Δlog likelihood=4.58, df=2, n.s.)

---

16 Coefficient p-values for variables in all logistic regressions are calculated using decrement to χ² (e.g., log likelihood) tests. Significance levels for constants are reported directly from the regression output.
These results are consistent with a view of explication as an intermediate variable, manipulated by the treatment, that increases the likelihood of noticing. Its effect was present in both subsamples and independent of participant background.

Rival Hypotheses

While the data associating the experimental treatment as a cause of the outcome is strong and while the analysis supports the explication's mediating role, the latter inference is subject to several rival hypotheses.

Recording Data

The treatment also is associated with an increased likelihood of recording data, leaving open questions about the role that data-recording plays. Table 3.11 shows the frequency of noticing summarized into the four categories of recording behavior.

<table>
<thead>
<tr>
<th>TABLE 3.11</th>
<th>Frequency of Noticing, Summarized by Recording Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Explicated</td>
</tr>
<tr>
<td>Recorded data</td>
<td>6 of 8 (75%)</td>
</tr>
<tr>
<td>Did not record data</td>
<td>3 of 3 (100%)</td>
</tr>
<tr>
<td>Total</td>
<td>9 of 11 (82%)</td>
</tr>
</tbody>
</table>

The probability of noticing varies considerably between the four groups—from 17% for those who recorded nothing to 100% of the three explicators who did not record data. At first glance, the table suggests a strong main effect of explication, a possible but weaker main effect of data-recording, and a possible negative interaction.

Table 3.12 offers the opportunity to compare the association between recording behaviors and noticing. A comparison on Models 1 and 3 shows that adding data-recording alone to a base model that contains explication does not improve model fit (Δlog likelihood = 0.51, df=1, n.s.). Adding both data-recording and its interaction with explication together (Model 4) also fails to improve model fit (Δlog likelihood = 3.54, df=2, n.s.).
TABLE 3.12
Logistic Regression Models Using Recording Behavior to Predict Noticing

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.16**</td>
<td>-1.59*</td>
<td>-2.37**</td>
<td>-2.76**</td>
</tr>
<tr>
<td>Subsample (control)</td>
<td>1.65*</td>
<td>1.13</td>
<td>1.71*</td>
<td>1.91*</td>
</tr>
<tr>
<td>Explication</td>
<td>3.25***</td>
<td></td>
<td>3.02***</td>
<td>10.54***</td>
</tr>
<tr>
<td>Data-Recording</td>
<td></td>
<td>1.44*</td>
<td>0.63</td>
<td>1.45</td>
</tr>
<tr>
<td>Expl. x Data-Rec.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2 Log Likelihood</td>
<td>39.54</td>
<td>49.56</td>
<td>39.03</td>
<td>36.00</td>
</tr>
<tr>
<td>df</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

*: p<0.05   **: p<0.01   ***: p<0.001

These models confirm the impressions one gets from the table as well as the increased likelihood of the second subsample to notice. Explication is a useful predictor of noticing and increases the likelihood of noticing whether or not one is recording data. The strength of the interaction term is the result of the combination of 100% noticing rate and small number of observations in the explicated/did-not-record cell. It is by no means certain (or necessarily even likely) that the interaction would increase in strength (and reach the 95% confidence level) if the sample were larger and there were more observations in that cell.

**Time Pressure**

Some have suggested that somehow urgency would play a role in the noticing process. Every attempt was made to control time pressure by ensuring that participants had enough time for the simulation and informing them of that fact.

During the interview, I asked people if they experienced any time pressure during the simulation and whether at any point they had felt they should speed up. Six responded that they had. Table 3.13 lists data on recording behavior, noticing and time spent for each. It should be recognized that this group is quite small relative to the sample making it unlikely, if not impossible, for it to have swayed the results one way or the other. Adding the time pressure variables to Model 2, Table 3.12 does not improve model fit ($\Delta$log likelihood = 0.03, df=1, n.s.).
TABLE 3.13
Participants Reporting Time Pressure

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Recorded data?</th>
<th>Explicated?</th>
<th>Noticed?</th>
<th>Time (minutes)</th>
<th>Time for similar rec. behavior?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>74</td>
<td>58</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>65</td>
<td>58</td>
</tr>
<tr>
<td>3</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>81</td>
<td>58</td>
</tr>
<tr>
<td>4</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>32</td>
<td>58</td>
</tr>
</tbody>
</table>

However, a qualitative look at these six is informative. Five of them took longer at the task than the average for people exhibiting similar recording behavior. Thus, while they may have reported time pressure, only one subject appears to have actually hurried. Of the ones in the treatment group, all recorded extensive data and explicated. Thus there is no evidence for qualitative change in their recording behavior, either. Of the two in the control group, one recorded no data, one recorded five time periods worth. While it is impossible to know how these people might have behaved in the absence of perceived time pressure, their behavior is certainly typical of others in their experimental groups.

Turning to noticing, one of the two members of the control group noticed, as did three of the four in the treatment group. These data, too, do not suggest any effect, positive or negative, of time pressure on noticing.

I am not suggesting that time pressure has no effect on the noticing process. I am simply pointing out that it cannot explain, or even help to explain, the treatment’s effect on noticing in this experiment.

*Speed of decision-making/Time to think*

A third alternative hypothesis is that the treatment does not result in a qualitative change in cognitive processing via explication, but instead simply slows down decision-making and thus provides more time for the processes that are already active to work. Since explication takes time and thus should be (and in fact is) correlated with playing time, the best test of this hypothesis would be an experiment using two interventions, one that induces explication and a second that slows participants down without otherwise
affecting them (e.g., without inducing explication, boring them, distracting them, etc.). Ideally, of course, such a test would use a repeated measures design and thus control individual differences, as well. Such an experiment remains a project for the future; however, there are data at hand to counter this rival hypothesis.

Were time, and not explication, the mediator responsible for noticing I would expect both data-recording and explication to be useful predictors of noticing since both take time. Table 3.14 reveals that explication and data-recording are comparably time consuming. Recording either data or hypotheses adds about 18-20 minutes to the average time for someone who records nothing, about an 80% increase. Recording both requires another 15-17 minutes, about a 40% increase. Both main effects are statistically significant (for explication p<0.05, for data-recording p<0.01); their interaction is not.

<table>
<thead>
<tr>
<th></th>
<th>Recorded data</th>
<th>Did not record</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicated</td>
<td>58.4 (16.5)</td>
<td>41.0 (6.9)</td>
<td>53.6 (16.3)</td>
</tr>
<tr>
<td>Did not expl.</td>
<td>43.0 (15.3)</td>
<td>23.4 (7.9)</td>
<td>28.5 (13.2)</td>
</tr>
<tr>
<td>Total</td>
<td>50.7 (17.9)</td>
<td>25.5 (9.5)</td>
<td>35.1 (17.9)</td>
</tr>
</tbody>
</table>

However, as we have seen, data-recording loses its ability to predict noticing as soon as explication is controlled. A quick look at Table 3.12, Models 1 and 3 reveals that data-recording adds essentially no explanatory power (Δlog likelihood=0.51, df=1, n.s.); the result cannot be attributed to insufficient statistical power. Thus, given these data the only way time could explain noticing would be if the activity of data-recording had some negative effect on noticing that was compensated for by its positive, time-induced effect. This explanation seems implausible.

In short, the data at hand offer no convincing evidence to conclude that time rather than explication explains the relationship between the treatment and noticing. In fact, they suggest the opposite.
**Motivation**

A final rival hypothesis is that motivation, not explication, is the variable that intervenes between the treatment and noticing. There are several ways to test this hypothesis.

It is reasonable to assume that highly motivated people would think more, consider more data, and analyze it more carefully. All this would take time. Thus one could use time spent as a proxy for motivation. There are counter-arguments, of course, for example that motivated people might work faster thereby compensating for the extra thinking that they do. Furthermore, individual differences between participants (e.g., their natural speed of thinking, the degree to which they are distracted, how sleepy they are, etc.) will introduce error into the measure and themselves may be associated with noticing. Despite the potential weakness of this measure, it seems reasonable to at least examine its relationship to noticing.

Model 1 in Table 3.15 shows that time is associated with noticing; people who take longer are also those more likely to notice. The addition of explication as a predictor in Model 3 improves model fit ($\Delta$log likelihood $=5.12, \text{df}=1, p<0.05$). Comparing Models 2 and 3 reveals that once explication is controlled time is no longer a useful predictor ($\Delta$log likelihood$=2.11, \text{df}=1$, n.s.). Models 4 and 5 indicate that when recording behavior is fully controlled, adding time does not improve model fit ($\Delta$log likelihood$=0.98$,

<table>
<thead>
<tr>
<th>TABLE 3.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression Models Using Recording Behavior and Time to Predict Noticing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-3.53**</td>
<td>-2.16**</td>
<td>-3.37**</td>
<td>-2.76*</td>
<td>-3.48**</td>
</tr>
<tr>
<td>Subsample (control)</td>
<td>1.04</td>
<td>1.65*</td>
<td>1.58~</td>
<td>1.91*</td>
<td>1.68~</td>
</tr>
<tr>
<td>Explication</td>
<td>3.25**</td>
<td>2.42*</td>
<td>10.54***</td>
<td>9.75*</td>
<td></td>
</tr>
<tr>
<td>Data-Recording</td>
<td>1.45</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expl. x Data-Rec.</td>
<td></td>
<td></td>
<td>-8.64~</td>
<td>-8.25</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>0.07**</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2 Log Likelihood</td>
<td>42.55</td>
<td>39.54</td>
<td>37.43</td>
<td>36.00</td>
<td>35.02</td>
</tr>
<tr>
<td>df</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

*: p<0.10  **: p<0.05  ***: p<0.001
df=1, n.s.). Furthermore, removing explication from Model 5 reduces model fit substantially (Δlog likelihood=6.59, df=1, p<0.05). In short, these data offer no evidence that using time as a predictor helps predict noticing. In addition, explication never loses its explanatory power, even when time is controlled.

We might also assume that recording behavior is itself a proxy for motivation. Since data-recording, unlike explication, is undertaken on the basis of individual initiative rather than formal request, it would make sense to assume that the people who recorded data spontaneously would be more motivated than those who did not. Of course, the analysis above (see Tables 3.11 and 3.12) really offers no basis for concluding that data-recorders are any more likely to notice than non-data-recorders. The interaction term in Table 3.12, Model 4 begs the question whether data-recorders in the control group are more likely to notice than their non-recording counterparts. Within the control group, 3 of the 7 data-recorders (43%) noticed compared to only 4 of the 22 non-recorders (18%). The difference, however, is not statistically significant (Fisher's exact one-tail test p=0.20178).

Furthermore, the data from the interviews do not support the conclusion that the data-recording members of the control group were more motivated than their non-data-recording equivalents. The following comments are drawn from interviews with non-data-recording members of the control group:

"It was frustrating because sometimes I'd be close, and then the next period I'd be off by 300 or 500. I kept trying... I thought, 'Oh, there's something I'm not getting here. There's something I missed.' It was frustrating and it was kind of fun." (4)\footnote{17}

"In general it was good fun." (21)

"I think I was motivated throughout the game... Sometimes I'd feel a surge of pride if I was close to zero. But then next time I'd be off by 1000 units." (28)

\footnote{17} Observation numbers are presented in parentheses so the reader may associate quotations that come from the same participant. In referring to participants, the pronouns he and she are selected at random and thus have no relationship to the actual gender of the study participant being discussed. They are not necessarily consistent for the same participant.
"I wanted to do well. In the middle, I got a little frustrated and thought about just putting in the suggested price. But I didn't do that. Toward the end it became more frustrating and challenging. I wanted to figure out what the hell was going on so I could do well. For me it became more of a challenge. I wasn't going to let it beat me." (31)

"I felt I was doing bad [sic]. It was challenging trying to figure out what was going on. Trying to figure out what was going on was fun." (34)

"My motivation was quite good. I had all these possible explanations. I wasn't trying to get finished as soon as possible. I was trying to figure out what was the problem... It was a challenge." (36)

"At the beginning I was just trying numbers. Toward the end of the game, there was more determination to get close and also more frustration if I was farther away. I'd been playing longer and should have known the trends. At the end I was definitely interested in continuing. It was fun. I'm kind of analytic and stubborn." (40)

"And then every five weeks or so, [I'd] go back through everything and make sure the model still holds up." (41)

There were five participants in this group reporting mixed or low motivation. The following is a list of every negative or mixed comment from the non-data-recorders in the control group:

"All the time I thought, 'There must be a way to beat this.' And I was embarrassed because I couldn't think of it... I thought, 'I'm probably not thinking about being creative and that's why I'm not beating it...' It's the end of the day, I'm tired. I tried a few strategies, but I felt there should be some nifty little trick to solve everything... If I were in the same room competing with other people I'd try harder probably because the price of embarrassment is more." (3)

"The first time you just play it and see what's happening. The next time you can go back and try more seriously... I think this is a silly game since I can't figure it out. There are things going on that are not apparent and not easily visible." (5)

"[I was motivated] up to a point. Then I figured, 'I can't [figure it out],'
about two-thirds of the way through the game." (7)

---

18 In weeks 15, 20 and 25, he spent more than double the average time used for his other decisions, a fact that lends credence to this statement.
"When the swings got bigger it was hard to stay motivated. I felt it was important to hit the mark." (12)

"It got more frustrating toward the end [but] I didn't give up." (37)

Listed below are typical comments about motivation from control group members who did record data spontaneously. They do not differ markedly from those of the first group, above.

"Toward the end it was getting irritating because of the variability and variables I maybe wasn't understanding... I was feeling pressure to do well each week." (13)

"It was frustrating. It made me curious. I would have liked to know how other people were doing... I wanted to excel." (23)

"My motivation didn't tail off. I thought I was winning." (30)

"When I'd be way off I'd think, 'Why not just pick the regression curve?' I did that a couple of times. I didn't abandon it. No, I gave up [and guessed] just a couple of times." (42)

I am not suggesting that motivation is unimportant for the noticing process. However, it cannot explain the treatment's effect on noticing, nor can it explain any potential-but-currently-undemonstrated differential rate of noticing between members of the control group who did and did not record data.

3.3.5. Sensitivity Analysis

Given the relatively small sample size and in particular the small number of observations in the explicated/did-not-record cell, it makes sense to examine the sensitivity of the results to outcome changes in this cell. Two questions drive this effort:

- How robust is the result that explication, rather than time and data-recording, explains noticing?
- How robust is the finding that explication does not interact with any other variables?

To answer these questions, I performed a series of tests, changing the outcome variable for one or two participants in the explicated/did-not-record cell. Only 3 participants fell into this cell; all of them were noticers (see Table
3.11). The small cell size means that changing one or two of its observations is likely to have a large effect; had any of these participants been a non-noticer rather than a noticer, explication would likely be a less powerful predictor. Such a change would also increase the explanatory power of time\(^{19}\) and data-recording.

Changing any one of these three raises the power of time as an explanatory variable. In two cases, changing a single observation meant that both time and explication became statistically significant predictors when the other was controlled (at the 90% confidence level). In the third case both variables remained statistically significant on their own; but neither was a useful predictor with the other controlled (even at the 90% confidence level). The same result obtained from changing any two of the three observations. In no case did data-recording or its interaction with explication become useful predictors.

In short, the main effect of explication is a robust result. The finding that time is not a useful predictor may be less so – which certainly makes intuitive sense – but does not call into question the role of explication. Finally, the finding that explication does not interact with other variables – time, data-recording, subsample, or background – is also a robust result.

I also used sensitivity analysis to determine the consequences, if any, of my decision to maintain the three non-noticers for whom the advertising variables were displayed continuously on the screen. In no case did dropping the three alter the results described in this chapter.

3.4. Summary and Conclusions

The experiment demonstrates a causal relationship between the treatment (directions to maintain a record of hypotheses) and the participants' likelihood of noticing of the advertising variables. Further analysis is

\(^{19}\) Participants in this cell took less time than the other noticers (on average, 41.0 minutes compared to 58.3 minutes for others; a difference slightly greater than the 16.3 minute standard deviation for all noticers). Had they been non-noticers, one might have attributed it to their playing the game more quickly than others. Given that they were noticers, however, time is not a helpful explanation.
consistent with the hypothesis that the treatment acts on noticing via explication as a mediating variable. Analysis of data-recording, time spent, time pressure, and motivation rules out the hypotheses that these variables, rather than explication, are the means through which the treatment has its effect. It also suggests that explication has its positive effect independent of whether people record data and how long they take to play the game. Sensitivity analysis demonstrates that explication's strong association with noticing is a robust finding. Based on these results, I conclude that explication does in fact make noticing more likely.

The treatment was effective in altering participants' behavior. People who were asked to record their hypotheses about the variables that might be useful predictors generally did so (11 of 13). This behavior did not occur spontaneously among members of the control group. This contrasts with data-recording (i.e., recording the values of variables), a behavior that emerged among participants in both treatment categories.

In this setting, with this task, and with this pool of participants, explication is a cause of noticing. Its benefits are greater than those of systematically recording data or of simply taking more time at the task. Before making any attempt to generalize, it makes sense to describe more fully the kinds of thought processes that noticers and non-noticers experienced. With a better understanding of some of these processes and how they are affected by explication, we can then begin to generalize beyond the specific circumstances of this experiment. Furthermore, I have argued that data-recording, time spent, and motivation cannot explain the results of this experiment. However, it seems unlikely that they would not somehow play a role in the noticing process. A more in-depth examination of participants' feelings and behavior during the experiment may offer some insight into the role that these variables might play.

In the next chapter, I present and interpret behavior of individual participants in order to describe more fully the ways in which explication has its effect. I conclude the chapter with a summary of the experimental results.
4. The Mechanisms that Drive the Explication-Noticing Relationship

4.1. Introduction

The previous chapter presented an experiment whose results link explication and noticing. The mechanisms that link the two have not been defined in any detail at this point. This chapter is a largely qualitative examination of the study participants' thought processes, based mostly on data from the interviews. It seeks to answer three questions:

- What are the barriers to noticing?
- What are the enablers of noticing?
- Do these barriers and enablers mediate the relationship between explication and noticing: are they the mechanisms through which explication has its effect?

The first two questions are exploratory and prompt an effort to distinguish variables and categories from the data in a process of grounded theory formation. The result is largely descriptive. The third question is empirical, a comparison of the frequencies of occurrence of these influences between those who explicated and those who did not. Thus the chapter is itself an effort to notice, to form a tentative model of the factors that contribute to noticing, and in particular to develop stronger hypotheses about the mechanisms through which explication may operate. Given the methodological limitations of the approach taken here, I think it is best to take this chapter as exploratory and its conclusions as hypotheses with some initial supporting data.

The next section (4.2) describes the data collection methods. Section 4.3, identifies and describes phenomena that, by encouraging, discouraging, or otherwise constraining the search process, function as barriers to or enablers of noticing. Three categories of barrier/enabler emerge from this inquiry: a category that affects the motivation to search (motivational barriers/enablers), another that directs the activities of the search process (script-based barriers/enablers), and those that constrain the content of what is considered (frame-based barriers/enablers). The section also proposes a model of the
search process exhibited by participants in the experimental study in order to clarify the role of the barriers and enablers.

Section 4.4 explores the relationship between explication and the barriers and enablers that emerged in the first, demonstrating quantitatively the role that the barriers/enablers play in the relationship between explication and noticing and developing further the relationships between explication and the barriers/enablers. It also demonstrates that the barriers and enablers defined here are only a partial explanation of explication’s effect on noticing; there are more barriers and enablers to be discovered.

4.2. Methods

The data for this chapter were collected primarily via the interview. However, in an effort to triangulate, these data were verified against the written protocols of participants in the treatment group, the records of their decisions and behavior recorded by the computer, and any notes that they took.

The interviews always began with a standard set of questions about types of help that might have been helpful to participants. The questions asked about the usefulness of help in making calculations, balancing tradeoffs between variables, testing the relationships between known variables, eliminating random noise, and identifying additional variables to think about. Participants were encouraged to verbalize their thoughts as they considered these questions. Their spontaneous comments at this point in the interview were often revealing; many are quoted in the sections below.

The interview then turned to open-ended discussion about:

- The variables that people considered and the strategies they tried during the simulation
- Their level of motivation and frustration and whether these changed during the game
- What they remembered from the case and what of that was important
- Whether they thought there was random noise in the system and if it affected them
- Whether the market changed or worked pretty much the same way throughout
• Whether they experienced any time pressure

• What parts of the screen were useful: I used an image of the screen on paper so that people could point and mark or circle things. If they had noticed the advertising I used a picture with the advertising window on it. If not, I used one without the advertising window because I did not want them to be affected by something that for them would be a new stimulus. By this point in the interview, participants who had used advertising had always mentioned it; so there was no ambiguity about which image I should use.

• What information they recorded and why

As needed, I asked more specific and closed ended questions within each of these topic areas. For example:

• Did you click on the advertising button? Did you think about advertising during the game? What did you see when you clicked on it?

• Did you ever give up and just guess for a while?

• Do you remember anything from the case or the instructions about advertising?

During the interview both the content of what was said and how it was said mattered. For example, some people exhibited surprise when questioned about advertising, lending credence to statements that they had not thought about it during the simulation. Sometimes comments were prompted by specific questions, other times they were completely spontaneous.

When people told me about the role of a particular variable, I would generally ask: "Was it something you thought of during the game and used, or is it something that you’re thinking about now?" Subjects were not vague in answering this question. Typically they would answer without hesitation and do so clearly, with either something like, "I thought about it during the game and tried to use it" or "It didn't occur to me until just now."

The sequencing of questions was very important in the interview, particularly those questions relating to advertising. Initial questions were open-ended, and no references were made to the advertising variables (e.g., "Tell me about the strategies that you used in the game?" "What variables did you think about during the game?" "Are there any other variables you considered that you have not told me about?"). Next I asked a series of
structured questions about variables that participants might not have told me about. I would preface it as follows: "I'm going to list a number of variables that you might have thought of. Please tell me if you thought about them, if you acted on them during the game, and if you think they might be important." Advertising was one among several plausibly important variables and was treated as no more or less important than the others, which included seasonal trends, business cycles, demand feedback from previous weeks, customer response to large price changes, important "price points" on the demand curve, and other kinds of non-linearities in demand. Later, in reviewing the picture of the screen, I would refer to each part, including the advertising button, and ask if they had been useful to the participant. Finally, the interview turned to the case and instructions. I asked what participants remembered from the case and instructions that was important or helpful. If they mentioned nothing about advertising, I would then ask whether the case or instructions discussed specific variables and finally ask them about advertising directly. Only at this point in the interview did I as interviewer differentiate between advertising and other plausibly important variables. The strategy was to avoid creating any demand effects associated with advertising and thereby minimize post hoc rationalization.

Not infrequently, participants mentioned advertising spontaneously or began discussing it in response to an open-ended question. In these cases, I followed up in the same manner as with any other variable they described. Depending on how it was referenced, I usually waited until the discussion of specific variables rather than single it out for special attention.

4.3. Barriers to and Enablers of Noticing

This section identifies potential barriers to and enablers of noticing based on participants' statements and behavior, and in so doing creates a partial model of the search process leading to noticing. While the discussion focuses on the noticing of the advertising variables, it is used to raise issues about noticing in general. At times the discussion includes variables other than advertising.

These barriers and enablers are presented as hypotheses. Participants' statements are reported verbatim and interpreted as though they represent
phenomena that enable or block noticing. A discussion later in the chapter addresses questions regarding the direction of causality.

The barriers and enablers fall into three broad categories, themselves emergent from the data. The first category includes factors that affect participants' motivation to search in an effort to improve their performance. The second includes scripts that determine what search activities are undertaken. The third consists of frames that focus search on certain types of variables and steer it away from others. Unlike scripts, which direct process, frames affect content/focus of search.

In Table 4.1, these three broad categories are disaggregated into more detailed categories. Each is framed as a barrier whose frequency of occurrence among noticers and non-noticers is reported. The subsections that follow each describe these detailed categories and present data indicating the role that each played in the experiment.

Four subsections follow, the first three corresponding to the broad categories of barrier. Each of these subsections presents data first and then closes with a summary for those who want to move through the data quickly. The fourth subsection is a discussion that presents a model of the inference process that helps show how each barrier or enabler has its effect.

There is, of course, an unspecified set of barriers and enablers that were controlled through the choice of the sample population, the nature of the task, and the experimental setting. These may be difficult or impossible to identify based on the data presented here. Thus the set of barriers and enablers listed here must be regarded as an incomplete set.

4.3.1. The Motivation to Search

This section discusses factors that led to the cessation of search or prevented it from being initiated in the first place. In the experiment, there were three ways in which participants' motivation to search for new variables failed. As indicated by Table 4.1, a very small number of participants seemed unmotivated by the task. More common were participants who gave up, concluding that search would not yield results (or further results, when participants felt they had learned something). A final group perceived no
### Table 4.1
Frequency of Barriers, Summarized by Noticing

<table>
<thead>
<tr>
<th>Type of Barrier</th>
<th>Noticed advertising (N=16)</th>
<th>Did not notice (N=26)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>freq</td>
<td>%</td>
</tr>
<tr>
<td><strong>Motivation-Based Barriers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Low task motivation</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>• Giving up</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>• No perceived need to improve</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td><strong>Any motivational barrier</strong></td>
<td>5(^1)</td>
<td>31</td>
</tr>
<tr>
<td><strong>Script-Based Barriers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Problems interpreting small N(^2)</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>• Drops vars when hyp. disconfirmed</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>• Restricted range of experimentation</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td><strong>Any script-based barrier</strong></td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td><strong>Script-Based Enabler</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Re-examination of own frames</td>
<td>5</td>
<td>31</td>
</tr>
<tr>
<td><strong>Frame-Based Barriers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Frames from the context</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>• Fixing the set of candidate variables</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>• Excluding what you can't control</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>• Frames constraining the form of data</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>• Specific (untested) interpretation of adv.</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Any frame-based barrier(^3)</strong></td>
<td>7</td>
<td>44</td>
</tr>
</tbody>
</table>

---

1 Since the totals reflect number of participants and some participants may exhibit multiple barriers, the "any barrier" line is not necessarily the sum of the disaggregated lines above. For example, one noticer exhibited two motivational barriers. Hence the total is 5, not 6.

2 This category includes several subcategories which are disaggregated and discussed in Section 4.3.2, below. One of them, looking at advertising only once or twice, is not included in these aggregate figures because it is necessarily confounded with noticing. It is discussed below and included when the data are tabulated by explication and by data-recording (with which it is not confounded) in Table 4.5.

3 This aggregated category does not include frames constraining the form of data or specific interpretations of advertising since they both are necessarily confounded with noticing. Both are discussed below and included in Table 4.5.
need to improve their performance and thus were unmotivated to search. The subsections that follow discuss each of these and their role in the noticing process.

**Task motivation**

With few exceptions, participants indicated that they were well motivated in performing the task. Task motivation is almost certainly an enabler of noticing. Only two participants provided unambiguous statements suggesting low motivation overall:

"I tried but I wasn't highly motivated. I got tired at the end. You can see, I stopped writing. I started thinking more about catching my bus. I became a little less patient... It wasn't a big concern to me whether I did well. For $200 or $300 dollars maybe I would have focused more." (15)

"The first time you just play it and see what's happening. The next time you can go back and try more seriously. I was thinking this is a silly game since I can't figure it out." (5)

In the latter half of the game, the first of these two used the suggested price for all but two of his decisions. The second participant's behavior was not consistent; in the latter half, he used the suggested price only twice and deviated from it both up and down in varying degrees suggesting some degree of experimentation. His suggested decision rule for others playing the game, however, was, "I'd say use the suggested price."

Two others offered mixed messages about motivation:

"It's the end of the day and I'm tired. I tried a few strategies, but I felt there should be some nifty little trick to solve everything... If I were in the same room competing with other people I'd try harder probably because the price of embarrassment is more." (3)

"When the swings got bigger it was hard to stay motivated. I felt it was important to hit the mark." (12)

---

Observation numbers are presented in parentheses so the reader may associate quotations that come from the same participant. In referring to participants, the pronouns he and she are selected at random and thus have no relationship to the actual gender of the study participant being discussed. They are not necessarily consistent for the same participant.
By far the most common response indicated a high level of motivation. Participants remarks about motivation (see Section 3.3.4 for more quotations) ranged from:

"In general it was good fun." (21)

to:

"I think I was obsessed with beating the computer." (24)

Overall, only 6 participants mentioned difficulty staying motivated. Given the sample (second-year MBA students) and the nature of the task (a management simulation game), it comes as no surprise that motivation was high. However, in other settings with a different population, there might not be such a high level of intrinsic interest in the task, a critical determinant of creative thinking (Amabile, 1988).

We can also see evidence beyond intrinsic motivation that affected participants. Of the 6 who showed a lack of task motivation, 5 were in the first subsample. As noted in Chapter 3, these participants were tested at the end of a semester when they were under great pressure preparing for exams, finishing papers, and worrying about jobs. Several participants made reference to outside pressures on them. Since subsample was a predictor of noticing (see Sections 3.3.2 and 3.3.4), the stress and distraction caused by factors unrelated to the task at hand may have limited the effort and attention that participants in the first subsample devoted to the task.

**Giving up: concluding the search process**

A total of nine participants exhibited signs of giving up. Several actually used the expression "giving up" during the interview.

"I think I really gave up around week 30." (3)

Others, however, expressed their feelings directly, thereby revealing what may have led to "giving up." One participant, for example, discussed his frustration during the simulation. When asked if he remained motivated despite it, he replied:

"Up to a point, yes. Then I figured, 'I can't [figure it out],' about two-thirds of the way through the game." (7)
This participant's comments and behavior during the experiment help clarify two points about giving up that are mirrored in other participants' experiences. The first is that giving up does not necessarily represent a failure in motivation per se. Rather it is an inference about the likely results of search. This inference can be an emotional response to frustration, a reasoned conclusion based on some near-economic rationality, or a response to problem framing. I discuss these points below.

People who reported giving up were not necessarily unmotivated, nor did giving up represent a complete failure of motivation. Both of the participants cited above expressed motivation initially, both in the way they described their experience and in the degree to which they experimented. The first reported:

"I thought, 'There must be a way to beat this.' And I was embarrassed because I couldn't think of it... I thought, 'I'm probably not thinking about being creative and that's why I'm not beating it.'" (3)

The second participant continued to use a decision logic that he had constructed earlier in the game and did so diligently. His behavior suggests that, while he was unmotivated to search for new variables, he was sufficiently motivated to continue concentrating and to perform as well as possible within the constraints of his decision model. In a sense, he, and many others like him, gave up the search for new variables but remained motivated enough to attempt to parameterize any models they had formed.

This pattern was particularly prominent among noticers. In their initial search efforts, they considered a range of potentially useful variables. Once they noticed the advertising variables only three resumed a search for more:

"Around week 20 or 25 I had a really bad week. So that's when I said there's another variable here. Maybe there is a saturation point!, maybe the demand curve isn't linear. So maybe there are other variables." (14)

"At some point in the game [week 29, according to her notes], I thought this is the year and this is Christmas. There should be some kind of seasonal pattern. Maybe this is Christmas here." (39)
The third person's notes confirm that he was concerned about a market saturation effect and that the market might "have a memory" and be sensitive to changes in price from week to week. All of the noticers put effort into estimating the effect size associated with advertising and four considered interactions or more complex relations between variables they had already identified. But only these three went beyond that to consider new variables.

As I have mentioned before, the participants in the experiment exhibited high levels of motivation. Thus it is not surprising that even among non-noticers, giving up due to frustration was somewhat rare. Only 6 participants in the entire sample gave up without having discovered some decision-making rules that they believed were effective in boosting their performance. However, there were other factors leading to giving up.

A few participants felt that while there might be some useful and undiscovered variables, they had neither the tools nor resources to identify them. Three, for example, mentioned variables for which they had no information:

[This participant mentioned variables to which he would have liked, but did not have, access. I asked him what they were.] "Competitor prices, who else is in the market. If I'm selling well I'd expect that others will enter the market. What are their prices?" (23)

In a sense, though, this participant had noticed these variables and simply lacked information to test his unstated hypotheses about them. Thus, the lack of tools did not interfere with noticing. However, the two others expressed the belief that they were observing and responding to all the variables to which they had access. Others noted more general types of resource constraints. For example, one participant claimed she did not give up but admitted that:

"With just three weeks to go I didn't think I could get it." (18)

One can certainly imagine a negative effect on search effort late in the game due to its fixed length. One can also imagine situations outside the lab where people fail to search, not because they anticipate search to be unproductive, but because they lack resources and time or because they expect the cost of search to outweigh its benefits. In a sense, giving up can be
economically rational; but ultimately it must be based on assumptions about expected benefit that are untestable.

Among these participants failure to search for new variables does not appear to have been a decision based solely on economic rationality. Rather, it appears in many cases to have been quite arbitrary and to be heavily influenced by framing. Let me offer an extreme example. One participant revealed:

"[I was] kind of bored at the end. I had had enough by week 30. I didn't put much into the last 10 weeks... It was very interesting initially. After I'd figured it out, my motivation waned." (22)

After noticing and incorporating the advertising variables into his decision logic, this participant did not make substantial efforts to go further. Giving up was not the result of a frustrated search effort, but rather a response to success. But how did this participant know that he had "figured it out?" He had learned something that had improved his performance; but what was the basis for his assumption that the game had no more variables to learn about or that he had solved the problem fully? A different participant with the same conclusion reported:

"I kind of gamed the system a little bit. I took into account how far you'd take the programming." (31)

This participant's comment makes clear the degree to which giving up is an inference and can be affected by one's framing of the situation. Two participants made assumptions about the level of complexity that I would be likely to program into a computer. In a sense, they were playing a different game, against me, than people who framed the task as a truer representation of a market (and thus potentially very complex).

Four participants' remarks conveyed an implicit framing of the task as a one-variable or one-thing-to-discover problem. They appear to have taken for granted a price-quantity relationship and then assumed that there was one level of complexity beyond that. Thus they expected to learn only one additional effect. Their perceived need for improvement was determined whether they had learned something:
"...there should be some nifty little trick to solve everything." (3)

"I did think there was something else to be looking at and it was a test of whether you'd get it. And I wasn't getting it." (18)

These examples may reveal a natural bias toward unidimensional framing, perhaps induced by the experimental situation. This "bias", of course, may be nothing more than an effort to simplify in the face of limited time, ambiguity, and limited motivation. As one noticer put it:

"I tried to figure out, you know, if you have more variables. More is better. But it's going to take too long. I think just the advertising was good enough." (25)

Whether the tendency toward unidimensional reasoning – that is, toward causal reasoning that assigns a single cause to each outcome – is a common bias remains an open question. Nonetheless, in the experiment, several participants' expectation of a unidimensional solution led to a cessation of search when, in fact, no inference that the problem had been solved completely could reasonably be drawn.

We can see that participants abandoned the search for new variables, sometimes because of frustration, occasionally because they simply felt they lacked the resources, cognitive or otherwise, to continue the search or to use its results, and not infrequently because of an unsupported inference about the task itself.

**The perceived need for improvement**

The previous section described participants' perceptions about the possibility of learning from additional search. This section focuses on the perceived need for improvement, rather than whether it is possible.

An important characteristic of the experimental task was that it provided accurate feedback, but no absolute standard by which it might be judged. Thus participants could tell if they were doing well or poorly only relative to their own past performance; they had no other reference. This aspect of the design, originally intended for purposes of realism, gave participants freedom in interpreting the simulator's feedback and, in particular, in inferring whether there was any need to improve performance.
There are two aspects of interpreting feedback that emerged as important from participants' behavior and comments, the specific variable used to measure and interpret the feedback stream and the participants' expectations with regard to the values that variable would take on. The participants focused either on the absolute magnitude of the difference between their target quantities and the actual number of orders received, on changes in this magnitude, or on the distribution of the deviations rather than their magnitude. The subsections that follow focus on these three variables that participants used and how participants formed and used expectations about the values for each in interpreting feedback.

The absolute magnitude of feedback

One group of participants seemed to acclimate themselves to a given level of inaccuracy (i.e., the difference between the number of orders received and their target each week):

"It was a new environment and I got used to it." (23)

"I really wanted to get that close [i.e., be accurate]. I learned what was a big mistake and not so big. So my anxiety decreased. In a sense, I was less ambitious." (34)

[Asked about giving up as a result of frustration with poor performance:] "I didn't give up. My goals changed." (37)

Without any clear indication that they could improve their performance substantially, each of these participants became used to the levels of variance that they were experiencing and treated them as ordinary and acceptable.

There were participants, however, who observed variation and immediately treated it as a disconfirming signal.

"A few hundred here and there is random noise but over a thousand seems to me too many orders to be random noise." (24)

"I felt helpless. I can't make things happen the way I want them to. I know that on the performance graph I should be along zero. But I was flipping around. (3)

"With only 2200 in stock, that [500 units] was 20% carried over. I just felt like that was too much... So there were obviously more variables."
Then I flipped through the guide to see if there were any variables." (14)

One participant, apparently without thinking, used magnitude of variation relative to the scale of the chart as an indicator:

[How well did you feel you were doing?] "I don't know but it [the chart showing differences between targets quantities and actual orders] was pretty close to the top of the scale and the bottom." [Did you notice that the graph was autoscaling?] (pause) "No." (8)

Both sets of interpretations – that is, those that confirm the acceptability of observed variation and those that disconfirm it – seem very, and equally, arbitrary. In all cases, the focus is on what is acceptable or how comfortable the participants feel, not on how small a magnitude of deviation is possible given the nature of the simulated system.

Changes in magnitude

Another group focused on changes in the magnitude of deviation in the feedback; they expected the size of the discrepancies to diminish over time as they learned. For these people, the failure to become more accurate over time was a source of frustration:

"I thought there would be a clear learning process, the noise to narrow down. I'd start narrowing then one would shoot off the screen. While I was playing I was a little frustrated. I expected it to narrow down." (11)

"I was getting frustrated with the things I was trying and I wasn't getting anywhere. So I thought, 'Where else can I go for help?" (8)

"You expect your performance to improve. It was a little bit frustrating." (32)

"I expected to see the high volatility on the graph getting smaller." (30)

A variation on the same theme were people who did not state explicitly an expectation about performance improving but who were alarmed when it seemed to get worse in the latter part of the game:

"I thought there was getting to be more variability, that it was getting harder to predict... My scores were getting toward 1000 at the end of the game. So something was going on but I wasn't sure what it was." (17)
It is possible that these people were focused on magnitude only, rather than change. But their focus shifted when performance appeared to deteriorate. For one participant, the deterioration was particularly salient because the scale on the performance graph, which adjusted automatically to accommodate any data it displayed, suddenly changed late in the game:

"When the scale was here and then all of a sudden the scale got bigger, it made a difference about how I felt about my performance." (10)

Participants expecting the deviation to become smaller over time sometimes offered theories to explain their failures to reduce or eliminate variation:

"The demand would shift and be random for a while. Then it would be still for a while... At some points there was lots of noise. It was random when the curve moved." (33)

"I expected to see high volatility on the graph getting smaller. I felt like I was learning. But the catch was the penalties got bigger and bigger... it just wasn't fair." (30)

In general, however, non-noticers whose performance did not improve over time felt that they had failed to solve the problem. They did not challenge their assumption that there was in fact something to be learned. The following typifies people's sentiments:

"I can't figure it out. There are things going on that are not apparent and not easily visible... There were a lot of other things going on that you didn't know." [Can you be more specific about that?] "You don't know what the external factors are that were driving demand." [Such as?] "The advertising, the weather, time, or I don't even know what I'm really selling so I don't know what external factors there are." (5)

"I felt a little helpless. I didn't have that much to work with... I kept asking, 'What else should I be looking at?' I did think there was something else to be looking at." (18)

"I felt I needed something else, more information." (23)

"I still feel there was something I was missing." (24)

Improvement was an aspect of performance that could be judged fairly unambiguously. For the most part, participants either noticed the advertising
and experienced substantial performance improvement or did not and experienced no change or a slight performance decrease halfway through the game. With the exceptions noted above, the non-noticers did not change their framing, that is, readjust their goals to match their experiences; they completed the game feeling they had missed the solution.

Aspects of feedback other than magnitude

Many participants recognized that they had no reference with regard to magnitude and were instead concerned with the mean of the error distribution rather than its magnitude:

"I started thinking about operations class and there being normal variation with processes and how when you tweak a machine to get the variation out of the process you end up making it worse. So I thought maybe the regression line would do better than I could. I’d just do what it told me and over time it would even out. When I followed the recommendation it seemed to oscillate pretty nicely around zero." (15)

This person’s focus was on the mean of the distribution. And, in fact, virtually all participants were concerned with whether they were pricing high or low generally. However, considering the mean of the distribution was certainly not a formula for success:

"I started out with $26 and that seemed just going along pretty stable so I used that as my base point." (16)

This person, rather than using suggested price as a starting point, used an estimate of the overall average, $26. After using that price for 6 turns and observing that the resulting distribution was centered on zero, he then noticed and began exploring the role of advertising. This person’s method of interpreting feedback provided no indication that he had forgotten an important variable – the target quantity (or, alternatively the suggested price).

Three participants were concerned with shape of the distribution as well as its mean:

"I was watching the data points develop into a line and decide [sic] whether there was a linear relationship or not and if it was linear was there a lot of noise or what. I was watching the epsilons [the difference between target and actual orders]. I put in the suggested price every
week... as that pattern developed – and it did very nicely – it became evident that the regression was giving me the best prediction. Everything else was beyond my control. By the 10th week it was pretty clear the relation was linear. I was looking for anything systematic like whether the suggested price was too low or too high, some kind of time series pattern. But it didn't look like anything was developing... at first it seemed like it [the performance graph] was bipolar and that concerned me. I would have expected more close to zero... Then I got a couple that were right on so I thought it's just by chance that they happened to land that way." (41)

This participant had substantial professional experience as a data analyst and clearly was quite observant. Despite the fact that she hypothesized for a time that the distribution was bimodal – the effect of the relatively powerful primetime/daytime advertising variable – she did not consider that another variable might be responsible. When asked, "Did you think that you were missing a variable, maybe one that has two values?" she responded:

"Oh (long pause). I see. No, I didn't ever think of that. Toward the end the spread may have gotten bigger, but it's tough to tell. I also got a few more in the middle which was really encouraging. So my guess about bipolar was off. On average it [the suggested price] was still giving me a good guess." (41)

This person's strategy could be summed up as, "I'm going to watch the residuals for evidence of a non-linear relationship between price and quantity ordered, for any shifts over time, or anything that would suggest that the straight line regression does not fit." The only indication that the regression was not a good fit, the non-normal distribution of the residuals, was not a strong enough signal for her to question her basic model.

"The question for me was how accurate is the model. If the model is relevant or significant then that's all that matters." (41)

Of course, the linear model was useful, the price a "significant" predictor of quantity ordered. However, the high accuracy of the model was completely unrelated to whether the unexplained variance could, in fact, be predicted. By insisting on statistical confirmation of her observations, she missed an opportunity to generate hypotheses and to notice variables that could improve her model.
Two others noted the bimodal nature of the distribution. Both were suspicious about it:

"I was kind of wondering after a while if there was some hysteresis... it did look like there was a line forming on the bottom and a line forming on the top. But I had no way of trying to analyze it to see if that was true. At least not easily." (5)

"The last thing I did at the very end it looked like I had two demand curves, a high one and a low one. And all my data for the last four weeks seemed to be on the low one. So the last price I chose was estimated from the low one." [What rule would you recommend for setting prices, then?] For me it would be very hard to say because I've thrown out my last mental model in Week 39 and I've gone with 'Let's use the lower demand curve or the higher.' If you look at the regression they give you, it goes down the middle. So you might as well (pause) I wasn't doing a whole lot better, so you might as well split the middle. Some weeks you'll be low. Some you'll be high." (6)

None of these three really developed a hypothesis to explain their observations about the shape of the distribution. Based on the ambiguity in the feedback, the first inferred that the distribution was not bipolar. The second chose not to follow up, in part because he could not easily analyze if there were two lines, and continued as though the distribution was not bipolar. Again the default response was to fail to explore a hunch. The third interpreted the feedback as disconfirming of his current model. Interestingly, he did not try to explain the presence of the two demand curves. It is not clear how he would have behaved had the simulation continued.

Summary

This section has provided examples of three sets of factors affecting whether participants did or did not seek to improve their performance in the game through search. The first was task motivation, determined first by participants' intrinsic interest in the task and then by the other tasks and pressures that competed for their attention. Low task motivation emerged only among the non-noticers and thus might represent a barrier to noticing. This comes as no surprise given that without it, search is unlikely.

The second was giving up, which can be thought of as an inference by participants that search would not yield useful results. Participants'
comments and actions suggest that while giving up did occur as a result of frustration, it also and most frequently depended on unsupported and, in fact, untestable inferences about what one could expect to learn in the game – for example, the assumption that there was only one thing to be discovered. For participants who framed the problem this way, the need to search was not a function of performance; it was determined by an inference about whether the search space was exhausted or not. From Table 4.1, we can see that although noticers did give up, they were perhaps less likely to do so than non-noticers; giving up is potentially a barrier to noticing.

The third factor was participants' inferences about the need to improve their performance. We can see that the participants used different feedback variables for interpreting performance. However, participants' perceptions of the need for improvement were shaped far less by their particular choice of variable than by their expectations about the values that these variables would take on. These expectations were typically unsupported, arbitrary, and left unchallenged during the game. The likelihood of this inference was almost as high for noticers as for non-noticers. It is important to recognize, however, that the noticers experienced a far higher degree of success from their efforts than the non-noticers and thus stopped searching later in the process.

There are several important messages to take away from this discussion. The first is that the motivation to search is not merely a function of intrinsic interest in the task. While intrinsic interest is important, it fuels a complex system where performance aspirations and motivation to search coexist in a series of three feedback loops:

1. Motivation to search drives search effort which in turn improves performance. Thus, aspirations are met and motivation to search reduced.

2. Poor performance, however, can erode goals, thereby reducing the motivation to search in the absence of any performance improvement.

3. Failed search effort may lead to frustration and giving up. The other forms of giving up can be addressed through this loop as well.
This system does not explain where the expectations come from initially, nor does it provide a mechanism for revising goals upward\(^5\). These goals, like task motivation, may act to fundamentally alter search behavior. At some level, expectations are necessarily arbitrary, the product of unsupported inference. There is simply no strictly rational way to decide that the unexplained variance in a system cannot be predicted or does not need to be. These assumptions, or problem frames, play an important role in the motivation to search and therefore can represent potentially critical barriers to noticing.

4.3.2. Scripts that Affect the Search Process

I have already discussed the most fundamental condition for search, that is, motivation. In the subsections that follow, I discuss a series of scripts that determine whether, and what, search activities are actually undertaken. Scripts are a type of schema that determine appropriate actions and thus specify processes. I include in the search process efforts to identify new variables, form new hypotheses, and determine their validity: the activities of model formation and use. Topics discussed below include:

- Learning based on small numbers of observations
- Scripts for responding to disconfirmed hypotheses
- Scripts restricting the range of experimentation
- Scripts for examining and responding to one's own frames and scripts

**Learning based on small numbers of observations**

The game was relatively short, a total of 40 time periods. Most participants responded to small numbers of observations and we can see that the use of small numbers was important, both for hypothesis generation and for testing.

*Hypothesis Generation*

In general, participants were quite adept at forming hypotheses, often on the basis of a single observation. I noted only one exception, a participant described above, who watched the performance (residuals) carefully, never

\(^5\) Actually, in such a system goals might be revised upward if, though chance, performance exceeded aspirations. This effect is a positive consequence of the second feedback loop.
deviating from the suggested price. She was waiting for clear feedback
disconfirming a simple linear price-quantity relationship. Early on, she
thought there was a bimodal distribution in the residuals – actually, the effect
of the advertising variables. She did not use that observation to drive
hypothesis testing. Instead she chose to wait for more observations in order
to do a test. Unfortunately, the pattern changed halfway through the game,
apparently disconfirming her bimodality hypothesis. From a hypothesis
testing standpoint, her caution might be commendable. But there can be little
doubt that her conservatism served as a barrier to noticing in the hypothesis
generation process. This participant, however, represents the only observed
"error" of this type, an isolated contrast to the other participants' spontaneity
in generating hypotheses.

Hypothesis Testing

In testing hypotheses as well, participants often made important
judgments rapidly, based on very small numbers of observations. As we
have seen in previous sections, these judgments, in turn, acted as constraints
on subsequent search. For example, in the second time period of the game
one participant began recording whether he priced high or low relative to
suggested price and whether he had too many or too few orders. He stopped
after only six time periods because:

"It seemed like pretty quickly there wasn't a clear pattern so I stopped
[recording] after six or seven periods." (12)

His conclusion that there was no clear pattern seems hasty, given that it
was based on only six data points. In a demonstration of the important and
odd decision-making that can follow such judgments, this participant stopped
recording data entirely. He thereby decreased dramatically the likelihood that
he would find patterns in the future (if his inference was not correct) and
eliminated the possibility of using historical information for subsequent
inference and hypothesis testing.

Another described his efforts to test one approach to pricing:

"[I tried] observing the oscillations and seeing if I could dampen them
by tweaking the price. I also experimented with if I wanted to do that if
there was some amount I could tweak the price. I didn't feel I learned
much." (15)
Examination of this person's actual pricing decisions indicates that the only time periods in which he might have tried this strategy were 6, 10, 11, 12, 21, and perhaps 35. This group of observations is hardly a reasonable sample from which to draw inferences. Another decided that demand was not a straight-line function of price after only a few trials. After describing an approach to the problem based on an assumption of non-linearity, I asked:

[How long did it take you to decide it was not linear?] "Four or five weeks." [Did you have a system?] "No, it was more like an art." (37)

This participant's conclusion that the price-quantity was non-linear had far-reaching consequences because it focused his attention throughout the entire game on the nature of the non-linearity, to the complete exclusion of other potentially relevant variables.

Perhaps the most extreme cases were participants who made important judgments based on single observations.

"I assumed weekend [demand] would be higher. But I tried it once but it didn't work out." (34)

In fact, his original hypothesis was incorrect and ultimately should have been disconfirmed; in the simulation, demand was lower when advertising took place on the weekends. The point, however, is that his process for testing the hypothesis left much to be desired. Another such participant reported:

"I picked it [the suggested price] once. It didn't work. So I started doing visualization [using visual estimates off the graph and ignoring suggested price]." (28)

There were many examples of participants who decided that a given strategy had failed after a very small number of trials. Unlike the participants described here, they were often correct in their conclusions (although such conclusions were unwarranted by the data). Whether correct or not, however, and seemingly independent of the sample on which they were based, these judgments were often binding.
Looking at advertising just once or twice

Hastiness was associated with more than just interpreting feedback. Seven participants in the study looked at advertising once or twice, but either distinguished no useful information or saw no change (although in each case the values of the variables actually did change). When looking at advertising, participants were implicitly testing whether the window contained useful information or whether it contained information that changed.

"I checked at the beginning and then two or three [weeks] into it to see if any numbers would have come up or anything changed. But since nothing did, I assumed that advertising was constant so price was the only variable." (31)

The advertising variables specified the timing of advertising on television. They were both represented as text, one indicating weekend/weekday time slots and the other primetime/daytime. Contrary to his claim, the values of these variables did change.

"I tried looking at the advertising cost information but it didn't let me look at anything. But then I remembered from the handout that you're not supposed to worry about advertising, so I didn't." (3)

[Did you click it once?] "No, I tried it twice." [And was the information different?] "There wasn't any [information]." [Was whatever showed up in the box the same?] "I assumed that it was. Maybe not. I'm not sure." (4)

[How many times did you click on it?] "I pressed it twice. It disappeared and I wanted it back again. Then it disappeared again." [Did it change?] "I don't know. I didn't even pay much attention to it." (6)

[Did you look at the advertising?] "Yes, I did at the beginning but it was blank and I didn't look at it again and I thought in the case it said not to worry about it." (10)

"I clicked on the button. It came weekly or something, not very important stuff. I looked at it only once. There were no numbers in the window: daily advertising or something like that." (p-7)6

---

6 In this chapter, I occasionally cite the comments of a participant in the pilot. I do not include the pilot sample in any tabulations or counts.
In the section titled "Frames specifying the form of relevant data", I discuss the reasons for these participants' failure to distinguish variables. Here, I wish to note how readily some participants were willing to dismiss the advertising function and their subsequent failure to reconsider that decision when searching for helpful variables.

**Scripts for responding to disconfirmed hypotheses**

I observed two participants who formed hypotheses about advertising, attempted to test them, and upon disconfirmation abandoned advertising as a variable without considering alternative hypotheses that included it. There was another participant who did the same with a different variable.

The first participant noticed advertising relatively late in the game:

"I realized I never looked at advertising because I started rereading the instructions. Yeah, there it is [referring to notes], week 27." (8)

He indicated that he tried to use it:

"It was pretty general. It didn't say anything in particular. It just said that this is a weekday and peak time or something... So you assume that has some impact or, being a marketer, I like to think that advertising has some impact on sales... I did use it." (8)

However, when he attempted to make decisions using it he points out:

"...sometimes the decisions I made didn't correlate with whether it was a weekday or a weekend... Then when I made decisions that way it didn't seem to work." (8)

His first statement is consistent with his behavior; based on the prices he quoted when he looked at advertising, he did not implement a program of pricing higher on weekends than during the week. Rather he was quite inconsistent. Examining the results, it is not clear how he could conclude one way or the other if his hypothesis was correct or not. What is particularly unfortunate is not that he concluded that a correct hypothesis was incorrect – an outcome easy to understand if testing the weekend/weekday hypothesis based on only a very few observations and failing to control for the powerful primetime/daytime effect – but that after disconfirming his hypothesis he chose not to examine any others involving advertising.
The second person's approach seemed even worse:

"I thought about looking at it [the advertising] during the game but I
got caught up in other things. I looked at it 4 times [actually 6]... I
assumed weekend would be higher. But I tried it once [and] it didn't
work out... I hypothesized and I tried it but it didn't work. So I didn't
bother to create a new hypothesis."

She looked at advertising in weeks 3, 14, 17, 18, 21 and 39. Her behavior in
weeks 18 and 21 is consistent with the description above. Both had weekend
advertising. In week 18 she priced 20 cents above suggested price and had
priced too high. Then in week 21 she priced 25 cents below and found that
she had priced too low. Given that the results for weeks 3 and 14, both using
weekend advertising, are consistent with the hypothesis that one should price
higher on weekends, it is not clear why she abandoned her hypothesis.
Nonetheless, she did and never looked at advertising again until the end of
the game.

Certainly a single or even two observations should not be sufficient to
disconfirm a hypothesis, particularly given the size of the variation in this
system relative to the price corrections this participant used. However this
judgment error was accompanied by a second and perhaps more fundamental
conceptual problem: dropping a variable entirely because a hypothesis
containing it was disconfirmed, the hypothesis-testing equivalent of throwing
the baby out with the bath water.

The test of a single hypothesis is a poor basis for eliminating a variable
from consideration. Yet these participants appear to have scripts specifying
exactly this behavior. It is impossible to know how many times this
happened with other participants who, for example, dropped time as a causal
variable because they could find no evidence for business cycles but never
considered that there might be some other kind of time trend or dependency.
Based on the difficulty of eliciting this information in the interview, I suspect
that this occurred more frequently in the sample than I was able to observe.

Given a combination of the small numbers bias and this judgment error,
the fact of noticing a variable does not necessarily make it likely that its true
role, if it has one at all, will be discovered.
Scripts restricting the range of experimentation

Depending on the quantity to be sold, the price suggested by the computer ranged from about $22 to $37. A quick look at the performance graph in Figure 3.1 suggests that unknown sources were producing routine variation of +/- 300-400 units in the number of orders. Later in the game, advertising's effect size increased, leading to even greater variance. By the end of the game, the typical non-noticer's performance graph has a standard deviation of around 600 units. Given the 30 cent to 100 unit relationship between price and number of orders cited in the case (the actual slope is closer to 35 cents per 100 units), one might expect participants to at least try experimenting with price changes of at least a dollar, perhaps even $2.00.

However, there was a group of 14 particularly conservative participants, 11 of whom priced within 30 cents of the suggested price throughout the entire game. The other 3 priced within 60 cents of it. The remaining participants in the study experimented frequently with prices more than $1.00 away from the suggested price.

This conservative group falls into three distinct categories:

- 1 participant did not adjust price at all.
- 8 "learned" from their very early experiences that they should not stray far from the suggested price.
- 5 experimented late in the game with the envelope within which they priced.

We have discussed the participant who did not adjust price earlier, in the section on feedback. Essentially, her strategy was to follow suggested price and watch the residuals appearing in the performance graph to decide if there was any reason to change strategy. She observed no feedback that suggested a need to adjust price from the suggested and thus did not do so.

The other participants were different, however. Based on comments in the interview and/or pricing behavior, 8 of those who stuck close to suggested price seem to have learned early in the game that deviating from suggested price was perilous:

"I was worried that it would diverge on me. But then I realized it wouldn't unless I did something really stupid..." [What do you mean
by stupid?] "Oh, set a price way off [from suggested price]. At the start I experimented with how sensitive it was. So I'd move the price 10 cents off. It looked like my method was going to work... I didn't want to go off more than +/- 30 cents." (42)

A second experimenter "found out that it didn't pay to vary your price from suggested price more than 2 or 3 cents. If you [varied it] more than that it didn't work well." (31) Typically these participants experimented in the first five time periods of the game and then settled into experimentation within close proximity of suggested price. The slowest in this group took until week 15.

It is not clear how the third group initially determined what was a reasonable amount to adjust the price. The only cue in the directions or case is the reference to the slope of the demand curve: "You'll see that we get about 100 additional orders for every 30 cents I lower price." When asked about the case, 14 participants mentioned the numeric reference to the slope of the demand curve, making it the second most commonly reported memory of the case, after the instruction "not to worry" about advertising which was remembered by 19. It is not clear how this cue could be linked logically to any particular constraint in the range of experimentation, it appears that some participants did in fact make this connection. Perhaps they were confused; alternately, this may just be an example of the kind of distant (and incorrect) links that the mind can make and then use automatically.

"I was going high and low. But not more than the 30 cents. She had mentioned the 30 cents changes for 100 [units]. I found that just 10 cents was hitting a couple hundred. A couple of times I ranged 30 cents just for fun. But most of the time it was 15 cents and half of that within 5 cents." (10)

Actually, this participant was within 15 cents of the suggested price for all but 6 decisions. His largest experiments were not at the beginning of the simulation, but in time periods 24-28. Three others followed this pattern, experimenting with broader deviations from suggested price about two-thirds of the way through the game. One concluded that the experiment did not improve performance and returned to working within only 2 or 3 cents of the suggested price.
The two remaining in this group were both noticers and they both reluctantly increased the size of the adjustments they made on the basis of advertising. Both resisted doing so. Although compensating for the advertising variables required adjustments as large as $2.50, one barely made it up to 20 cents, the other up to 60.

"I was conservative. I'd try pricing up 8 cents in primetime. There was more demand. I put it up 10 and then next time 12 and up to 15." [Did you try larger amounts?] "No. I anticipated terrible results. I was just trying to get a handle on what I had and I didn't want to lose any data points." (29)

"I kept [deviating from suggested price] more and more and it still wasn't enough." (20)

Both of these participants kept experimenting with larger and larger price adjustments. But expressed surprise or doubt about increasing the size of the adjustments they made, even as they saw in the feedback the need to do so. As a rule, participants were conservative:

"It took a while to get the range [for price adjustments from suggested price]... I was conservative, 20 cents or 30 cents. Finally toward the end it was like a buck." (30)

The diaries of noticers in the treatment group were littered with comments such as:

"Primetime ads give greater price premiums than expected." (9)

"Prime weekday. Try + 75 cents. Not enough!!" (Later) "Primetime weekday again. Saturated from last week? +100 cents. Wow! Not enough." (A few turns later) "Prime w'day + 1.15! Still not enough!" (And, again, later) "Prime weekday, +1.30! Still not enough!!" (19)

At the end of the game, this participant was making adjustments as large as +/- $1.80. Compensating fully for advertising would have required making adjustment of +/- $2.50. In this case, trial and error learning won out; conservatism merely slowed the process. But in the cases mentioned above, participants restricted the range of their experimentation and did not appear to push the envelope.
There are several potential explanations for this phenomenon. It is clear that these participants relied on directional cues from trial and error learning and, at best, estimated only informally the size of the price adjustments they needed to make. Within this group, none calculated the price offset appropriate to compensate for the quantity differences they observed, thereby demonstrating the insufficient adjustment that accompanies an anchoring and adjustment heuristic (Tversky & Kahneman, 1974, 1987).

The informal estimation of the appropriate adjustment size allowed other factors to influence decision-making. For example, participants' motivation to perform well led to loss aversion. A number of the participants cited above mentioned either a fear of ruining an otherwise good performance by "doing something really stupid" (42), that is by setting prices that in their estimation deviated substantially from the price suggested. One mentioned:

"I was feeling pressure to do well each week." (13)

Others imagined that more radical experimentation would not yield useful information; they expressed a fear of "losing data points" (29). Thus, participants' conservatism may have reflected one of two particular rationalities, perhaps analogous to the suppression of experimentation environments where the performance risks associated with mistakes are great, and a second where the expected utility of anything other than conservative experimentation is viewed as low. Alternatively, one can frame the judgment task as one of adjustment from an anchor, the suggested price, and thus subject to the natural human bias toward insufficient adjustment demonstrated by Tversky and Kahneman (1974).

Scripts for examining and responding to one's own frames and scripts

There was a group of participants who exhibited skepticism about all their conclusions. I observed participants skeptical about their interpretations of the instructions, evaluations of the feedback they received, and assumptions they had made about the advertising variables.

When participants read instructions, many interpreted them to indicate that the advertising was not useful. Several exhibited skepticism:
"She said, 'Don't worry about advertising.' I didn't know if it was a trick for me not to look too soon." (20)

"It said don't worry about advertising. But I figured it's on the screen so I should look at it... I realized, 'Why's it there? It must be there for a reason.'" (8)

Most participants read the instructions and the case only once. Only three reported reviewing them, typically because they were frustrated:

"I was getting pretty frustrated with the things I was trying and I wasn't getting anywhere. So I thought, 'Where else can I go for help?' Reread the instructions. I reread through the things that whoever said to look at." (8)

In all cases, the people who reread the case at least looked at the advertising window. They seem to have approached the directions with an open mind, in fact, with the attitude that they did not understand the problem, and each found new things to try, including the advertising.

When asked how well they had performed, most participants claimed to have done poorly. There were only few – four of the five were noticers – who responded that they had no idea. They had not gotten any feedback that would allow them to answer the question and were explicit first in acknowledging the ambiguity in the feedback and second in assuming that there was more for them to learn. One even framed the interpretation of feedback as a choice:

"What's noise and what's not? It's hard to say... So you can say it's either random or there's a lot more going on... but it was for me more intriguing." (6)

These participants recognized the ambiguity inherent in the feedback they received and therefore used it as an indicator of any learning on their part but not as an indicator of how much more might be possible.

The ability to draw conclusions and yet remain skeptical, to make decisions on the basis of tentative knowledge and yet remain open to new hypotheses, seems important for participants' success at this task. Successful participants seem to have questioned their assumptions, even basic

- 118 -
assumptions such as the apparent lack of usefulness of the advertising window.

These participants may have had scripts that helped them maintain conclusions as open issues longer than others or that allowed them to maintain multiple hypotheses simultaneously. They appear to have had an instinctive skeptical response to their own knowledge and to have asked themselves what other conclusions they might have drawn from the same information. Even when aware of their tendencies to close down, however, participants still may have used inappropriate scripts:

"I felt guilty not looking for the outside information on advertising because I know from previous employment that I have this habit of using the tools in front of me rather than trying to expand my view and think about what things could be affecting it." [You mean you feel that way now?] "No, I felt guilty during the game." (12)

Another commented that in complex situations in the past he has focused on too few variables: "My bandwidth [i.e., the number of variables he'll consider simultaneously] tends to become zippity [i.e., small]." Despite this self-knowledge and his awareness of its consequences, he did not try searching for more variables.

The enabler of noticing that emerges is a combination of awareness of one's scripts (and frames as described in the next section) and a healthy skepticism about them. Both appear required for participants to reconsider their ideas and choices of action during the simulation.

Summary

This section has presented data suggesting that participants' search behavior followed patterns that, at times, interfered with noticing. Participants' readiness to respond to small numbers of observations, so critical for generating hypotheses, became a problem when it also served as the basis for rigidly held conclusions from hypothesis testing. Some participants made a serious judgment error when, on disconfirmation of a particular hypothesis, they abandoned not only the hypothesis but the predictor variable that it included, as well. Others fell into a pattern of experimenting within a narrow price range (that is, a small deviation from
the suggested price), a practice that slowed learning (in the context of this task). Finally, many participants did not turn their inquiry to their own assumptions and search for intellectual constraints that they may have created for themselves.

Although the barriers described here each appear to have affected individual participants' behavior, Table 4.1 does not offer data supporting relationships between all of them and noticing. Rather it suggests that noticers were more likely to reexamine their own frames, less likely to work within a restricted range of experimentation, and perhaps less likely to make judgment errors based on very small numbers of observations.

4.3.3. Frames that Affect the Direction of Search

This section reports on problem frames that direct search. The scripts in the prior section directed activities in the search process. Frames have their effect by controlling content; they affect what people search for in their conscious efforts and what is perceived through automatic cognitive processes.

Like the scripts discussed in the previous section, the frames presented here can be viewed as inferences that are sometimes based on little or no data. These frames guide the search process, specifying content-related boundaries that limit the territory that is searched. Although essentially untested, these frames function as important constraints on problem-analysis and the search for solutions.

I observed five types of frames directing search among the participants in the experiment. All but the last are independent of the nature of what was to be found (the advertising variables).

- Frames supplied by the context that exclude variables
- Frames that fix the set of candidate variables
- Frames excluding what you cannot control
- Frames specifying the form of relevant data
- Frames based on specific interpretations of advertising
Frames supplied by the context

It makes sense that participants would be influenced by the case and instructions. As people tackle an unfamiliar problem, they are likely to draw information about it from any available source. Both the instructions and the case refer to the "advertising cost information." The instructions mention the button and describe the information that it reveals, including the values that the two variables can take on. The references in the case were intended to establish the advertising variables as an unobtrusive, not-necessarily-relevant part of the task environment, thereby making them something to be discovered. The "boss" in the case makes reference to the advertising information by saying, "The spreadsheet also helps me out in calculating advertising costs," but then draws attention away from them by saying that, "But I don't want you to worry about costs yet. You should focus on accuracy."

A large portion of the sample (17 of 42) remembered references to the advertising from the case and claim to have interpreted them as directions to not use the advertising button. Some participants appear to have framed the problem in a limited way based on the instructions, eliminating advertising from the set of candidate variables before they even understood what it really meant.

"The instructions specifically said, at least my interpretation was, 'You're not supposed to worry about advertising as part of your decision. Just think about price.' I realized halfway through that it wasn't just the price, but by then I was so frustrated that it didn't occur to me to go look at advertising." [Why not click on it?] "I can't give you a good reason why other than I read it and I quickly forgot about it." (24)

[Spontaneously, when asked about any other variables considered,] "I never looked at advertising costs. It said not to worry about it. So I didn't worry about it." [Did you think about it during the game?] "No. I didn't even spot it until right now." (37)

[In discussing the screen: Did you click on either of the buttons?] "No. I didn't want to quit and this (he points to the advertising button) was just extra information that the case didn't suggest you need to use." (33)

[As part of a list of questions about specific variables - this is the first mention of advertising in this interview: Did you think about
advertising?] "I wasn't supposed to mess with advertising, was I? I wasn't paying any attention to it." (15)

Each of these participants was aware of the advertising button by the time of the interview and all spontaneously and immediately recalled the case and their interpretations of it. This last participant and I moved on to discuss other variables, many of which she had not thought about, either7. When asked later in the interview what she remembered from the case, the first thing she said was:

"Advertising was something you could play with but it's not something that I did play with. For some reason I wasn't supposed to mess with that. I think that's what the woman in the case indicated." (15)

Others did not mention the case and could not remember any references to advertising when asked directly. However, they somehow had the impression that advertising was not useful.

[Were there other variables you thought about?] "There wasn't enough for me work with. All I saw on the screen was price and quantity and felt as if I should be looking at other things but I had no clue what I could be looking at." [Then, discussing specific variables: Did you think about advertising?] "I didn't think I could access anything about advertising. I didn't think it would affect it, to be honest." (14)

[Did you think about any other variables?] "I thought about advertising cost information. I thought if it would have been helpful to have it." [And did you click on the button?] "No. I didn't ask about it. I kind of figured if I was supposed to do it (long pause). That's a good question. Why didn't I do that? But I didn't... Was I supposed to use it? Anyway, I did think that it might be helpful. Maybe if I'd had the impression I was allowed to use it I would have. I definitely am very much [a person who thinks you should] get all the information you can get your hands on to do the analysis. I definitely would have if I could." (36)

Where, if not from the case, did these participants get the impression that they could not access anything about advertising?

---

7 None of the participants, noticer and non-noticer alike, considered all or even most of the variables in my list; so this was not atypical or a sign of poor performance or motivation.
There was a small group of participants who looked (or tried to look) at the advertising and failed to distinguish any variables. I will discuss failure to distinguish in a subsequent section. Here I wish to focus on their descriptions because in describing events and their interpretation of the advertising information, the case plays a prominent role:

"All the time I thought there must be a way to beat this and I was embarrassed because I couldn't think of it... [In discussing the parts of the screen] I tried looking at the advertising cost information but it didn't let me look at anything. But then I remembered from the handout that you're not supposed to worry about advertising, so I didn't." (3)

This participant tried twice to click on the button part way through entering a number (i.e., he put in several digits and tried clicking during the process of data entry). Excel does not permit pointing-and-clicking with the mouse while entering data into a field (i.e., after you have entered one or more digits but not completed entering the information). As a result, his effort was met with a beep and no visual results. Clearly this participant had not entirely ruled out advertising as potentially useful information. However, the combination of his inability to click on the button combined with his memory of the case led to a framing of advertising as irrelevant. There were others who clicked on advertising and appear to have been influenced by the case:

"I looked at the instructions again. It said something about advertising that made it sound like it wasn't useful. So for the first 10 rounds I didn't look at it. Then I asked you, 'Can I click on this advertising cost info? Is it useful for me?' And you said, 'Oh, sure.' I just looked [for] numbers, didn't see any, and concluded that it's probably useless." (4)

---

8 This dual modality (enter data or point and click) is a feature common to most spreadsheets and all dialogue boxes in window-based programs. Had he tried clicking on the advertising button at any other point in the process the information would have been revealed. No other participant experienced this difficulty; all who tried clicking on the button saw the advertising window.

9 Actually I said, "Do whatever makes sense given what you read in the case and instructions."
[Did you look at the advertising?] "Yes, I did at the beginning but it was blank and I didn't look at it again and I thought in the case it said not to worry about it." (10)

If we are to take these descriptions at face value, each of these participants already had an impression from the instructions that advertising was not important. They all tried to look at advertising, two successfully, and, when they failed to distinguish useful variables, they used interpretations of the instructions as part of the basis for framing the problem.

Clearly, all the study participants did not interpret the case in this manner. Many looked at the advertising and used it. Of them, three remembered the references to advertising:

"She said, 'Don't worry about advertising.' I didn't know if it was a trick for me not to look too soon." (20)

"Don't worry about advertising. I thought this was only the first run of the simulation but I just decided I'd look at it a couple of times." (22)

"It said don't worry about advertising. But I figured it's on the screen so I should look at it." [Later, when asked about advertising in the case:] "It said it wasn't relevant or not to worry about costs. So maybe in my mind I thought advertising costs. And then I realized, 'Why's it there? It must be there for a reason.'" (8)

These participants were aware of and questioned their interpretation of the case (e.g., an interpretation that identifies advertising as unimportant or otherwise outside the scope of the problem). The first two never really accepted a specific interpretation of the case, while the third, frustrated late in the game, revisited the case and instructions and reframed the problem at that point. Thus we can see two ways in which people challenged an interpretation of the case, as skeptics from the start and, driven by frustration, as information seekers, looking for clues that they might have missed the first time through. Overall, however, we can see how cues from the context influenced them quite easily.

The fact that this was an experiment served as context, leading some participants to assume that there definitely was something to be learned. While these frames played a role in motivating participants, I discovered none that directed search as did the inferences drawn from the instructions.
Frames that fix the set of candidate variables

The previous subsection presents frames that participants inferred from the context of the task that excluded advertising from the set of potentially relevant variables. But study participants exhibited a different, more general approach to constraining this set. At some point, they appear to have simply limited themselves to the variables they had already identified.

"I made the assumption that there's a whole lot of variables out there that are probably affecting it in a certain way. But I'm just going to focus on the last price to give me a general feel to get me going in the right direction." [Later, when asked whether he considered advertising:] "I didn't know if advertising would affect sales this week, 4 weeks out, or 8 weeks out. There was too much complexity there. I just didn't want to think about it." (p-9)

"I didn't know if it would give me any trends and it would have added more variables and I already didn't know what the variables I had were doing. So if I'd had more variables I don't understand I wouldn't have gotten more out of it." (5)

"I thought each part of the line might have different characteristics... I intended to keep track of that but tracking my four variables [in this case, price adjustments associated with each type of advertising slot] consumed all my time." (29)

"When uncertainty or ambiguity is high, I get focused on a very small set of things I can deal with. My bandwidth is zippity. [Then at the very end of the interview: suppose you had remembered about the advertising from the instructions, would you have clicked on it?] I don't think so because of the bandwidth issue." (13)

Each of the participants chose to constrain the set of variables early in the game. They are each acting, at least in part, out of a desire to limit complexity. However, restricting the set of variables did not always lead to simple models. Often the models that participants generated became quite complex in other ways. For example, one participant described a model in which the demand curve would shift up or down periodically. These transitions were marked by periods where there was lots of random variation. After a shift, the noise would attenuate and then the curve would start to shift again. Another hypothesized a curve with a very irregular and "spiky" shape and attempted to plot the wildly non-linear shape of the demand. Another developed a theory involving distinct regions of the curve where the amount of
randomness varied, where there were discontinuities, and where the curve had fundamentally different shapes. The complexity of people's theories did not appear to stimulate the question, "Could this all be explained by the presence of another variable?"

Noticers, as well as non-noticers, appeared concerned with limiting complexity by restricting their universe of variables. When asked if she had considered any variables in addition to those she had already mentioned (quantity and advertising), one participant responded:

"Not really. It was still jumping around so much. I didn't want to expand the strategy." (26)

Another cited a combination of avoiding complexity (or not putting in the effort necessary to handle complexity) and satisfaction with his progress thus far:

"I tried to figure out, you know, if you have more variables. More is better. But it's going to take too long. I think just the advertising was good enough." (35)

Constraining the set of variables was also a practical response to lack of information. Why consider a variable about which one cannot get information?

[As part of his description of how his strategy evolved:] "Operating in a vacuum, I didn't think about what other things could be affecting it. But I did think that possibly other things could be affecting it. I didn't take that on to figure out what those things might be. I'm somebody who uses the tools I have to solve the problem rather than looking for alternative ways of looking at the problem." (3)

[As part of his description of how his strategy evolved:] "Once I knew that I wasn't going to have any information on other things going on I didn't even think about it [figuring out other variables] 'cause I couldn't deal with those [i.e., wouldn't have the data to respond to them]." (31)

These participants did not choose to limit the set of causal variables; they saw it as fixed by the information to which they had access. In a sense, each of these participants "gave up" on searching within the domain of new variables.
Another factor that constrained the variables that participants considered was their prior recording behavior. Participants did think of specific, potentially useful variables late in the game, but without having recorded their values they could draw no conclusions. Interestingly, they often failed to start recording them, even when they were less than halfway through the game.

[In discussing why he didn't use and record data on a specific variable:] "I thought about it, but I didn't [use it]. It took me between 5 and 10 weeks to figure out that was information I wanted. By then I thought, 'I already don't have that information.' So I just continued, thinking I'd remember. Then, by week 20, I realized I wasn't remembering and then I was already halfway through." (40)

Had the game been an ongoing activity rather than a task of fixed length, this participant and others like him might have started recording and using the variable in question. However, one can also imagine how in organizations framing could become embodied in, and virtually enforced by, systems for collecting, organizing, and presenting information and by standardized approaches to decision-making.

Frames maintained by data-recording

Some participants who recorded the values of variables concluded that they were not relevant and stopped recording them. Typically noticers recorded the value of the advertising variables, the price they set, the suggested price, the target quantity and the quantity actually ordered for a period of time. On becoming more confident in a hypothesis, several of them changed their recording scheme to one that made recording information about each of the four types of advertising simple, but losing track of the time sequence of data. They also might stop recording other useful information; for example, they might record performance rather than both performance and the actual number of orders. In these situations, they restricted the data they recorded to that which was necessary to test their hypotheses. When the hypotheses proved incorrect, as most that did not involve the advertising variables did, these participants were limited by the data they had.

One can imagine how busy people would tend to avoid creating complexity for themselves and, where they lacked tools or felt they did not
have access to information, would simply accept constrained framings of the problems they face. One can also imagine that people would create tools based on their current framing that would limit their view and thus the potential frames they might adopt in the future. These participants also demonstrate that modeling with fewer variables does not guarantee lower complexity; in fact, in situations where the set of candidate causal variables is missing one that is important, the result may be a complex, unwieldy model.

_Framings that exclude advertising temporarily_

A group of people framed advertising as irrelevant for the time being, a matter to be taken up at a later date. It is easy to see how they might have drawn such an interpretation from the case:

[What do you remember from the case?] "The case did say not to worry about advertising. You know, you play it once and then you start playing with everything." (5)

[What do you remember from the case?] "She [the boss in the case] said advertising cost was something I shouldn't worry about, to get a feeling for the model first. And I did that." (13)

[Did the case say anything about advertising?] "I remember you could push a button for advertising costs. This is only the revenue side. There's also cost issues. Some of those variables could definitely influence my order strategy. But I wasn't using them. I assumed at a later time we'd come to cost and we'd deal with those issues. At this point I was just trying to make the demand side work for me." (p-9)

[What do you remember from the case?] "They said they would go over the advertising costs but it's not important right now." (40)

Temporarily framing variables out of a problem makes sense from the perspective of managing problem complexity. One might argue that it is rational in terms of cognitive effort to exclude variables temporarily. What is striking, however, is that even in the face of frustration these temporary frames go unchallenged. Such persistence makes me question whether there is a tendency for these frames to become permanent.

_Frames excluding variables that you cannot control_

One repeated theme was the idea that things that cannot be controlled are somehow irrelevant.
[Did you click on advertising?] "I didn't think of advertising because I didn't seem to have any control over it. So I considered my decisions to be made strictly on the basis of the information that I had. I put it out of my mind. I didn't [press the button]." (3)

Given that the instructions make no mention of setting or adjusting advertising, it is perfectly reasonable to assume that one cannot alter it. What is not clear is why one would assume that it was not relevant. Nonetheless, control (along with other factors that I discuss below) appears as a factor in several participants' descriptions of their thinking.

[Did you think about advertising?] "I clicked on it once. This box came up saying that where advertised [sic]. Since I couldn't play with it, I just ignored it. (long pause) Oh, did it change over time? Once I couldn't adjust it, I (pause) Oh, that's true, it might have changed." (7)

[What other variables did you think about?] "There wasn't access to the advertising stuff so that the only variables I saw were price – that we had control over." (10)

"The instructions say that I don't have any control over it. That's why I didn't look at it until week 5. Once I saw that it mattered, it bothered me that I didn't have control over it." (19)

I propose that participants inadvertently frame the task as a control problem, a reasonable framing given the instructions, but one that directs their attention (inappropriately, in this case) away from variables over which they have no control. Were they to apply a prediction-frame, all variables would take on equal relevance, independent of the degree to which they can be controlled.

**Frames constraining the form of relevant data**

The advertising window is shown in Figure 3.1 and consists of textual variables color-coded as data. There were 6 people who clicked on the advertising and never used the advertising variables. Of them, 4 appear not to have distinguished variables in the window:

"When I clicked on it I didn't get any information. So, it was kind of blank. When I clicked on it it said, 'this week,' and nothing. I don't know if that meant - and I didn't try it - that I could put in data in there." [what do you mean, blank?] "I didn't get any numbers. It didn't really help me." [What do you mean, numbers? Costs?] "Not just
costs. Something like how many times you advertised. And then if you did that over time you'd be able to see you know airtime, daytime, 3 or 4 times this week and you could look at demand and say 3 times means demand goes up but if you only did it once that didn't really change anything. So you can see the relationship between advertising and demand..." [Did you click it once?] "No, I tried it twice." [And was the information different?] "There wasn't any [information]." [Oh. Well, was whatever showed up in the box the same?] "I assumed that it was. Maybe not. I'm not sure." (4)

[Did you look at advertising?] "Yes. I did at the beginning, but it was blank and I didn't look at it again." [When you say blank, what do you mean?] "There was no data there. I didn't look back at it again. Because when I started up I looked at it but there were no numbers. It had some categories but I got kind of engrossed in other things." (10)

"I tried it [the advertising button] and I said, 'There are no numbers here.' So I went off and played another period and so then you go see if there are any numbers in it now, and there aren't. I just looked at the numbers, didn't see any, and concluded that it [the advertising window] is probably useless." (8)

These participants had expectations about the nature of data, namely that it would be numerical. It is quite possible that two participants within this group actually distinguished text variables – they both remembered some aspects of the information in the window – but discounted them because they were not numeric.

One of the participants who looked at advertising twice appears to have also been expecting numbers and failed to distinguished any variables. Contrary to what he says, the information displayed in the window was different in the two time periods.

"I checked out the advertising cost information and it seemed to apply. [I saw] a table with nothing in it..." [Did you check more than once?] "I checked twice. I checked at the beginning and then two or three [weeks] into it to see if any numbers would have come up or anything changed. But since nothing did, I assumed that advertising was constant so price was the only variable." (31)

Einhorn and Hogarth (1982) argue that people expect causes and effects to appear similar. Given a numeric effect, participant may adopt a frame that tends to screen out non-numeric information; since the advertising variables
are in a text form, they are filtered out. Alternatively, the only-numbers-are-relevant frame may be the product of the sample's quantitative background and training at Sloan.

Several participants' comments about seasonality point to another type of failure to distinguish potentially useful variables. When asked whether they thought about seasonality as an influence on demand, they responded:

"I wasn't given any information on this so I didn't even bother to think about it." (30)

"I didn't see where to get this information. It just crossed my mind and I didn't readily see how to assess whether the season was changing." (38)

"I tried to ignore it. There wasn't enough information." (34)

"If there were a calendar there you would have known to think about it [seasonality]. The fact that it wasn't there meant that you couldn't think about it." (p-6)

The time scale was in weeks and performance was plotted chronologically. Certainly well into the game, one might have developed hypotheses about the correspondence between the weeks indicated and the calendar year. In this case, inference and translation were required to get from the information provided, week number, to information in a useful form, calendar week. Only one participant in the study actually made such an inference and used season to help predict demand.

**Frames based on specific interpretations of advertising**

The frames discussed up to this point have referred to the generic process by which frames are formed and used. Up to this point, the discussion has been separable from the specific object of noticing (i.e., advertising). However, there are a class of observed frames that cannot be generalized; they are specific to advertising. It appears that some study participants made assumptions about the content or meaning of the advertising variables without even looking at them:

[Did you click on the button?] "I didn't. Advertising cost information to me was how much it would cost but not necessarily how much they were spending. So I made an assumption and that may have been a
mistake. If I had thought that was giving me information on how much we were spending on advertising I would have looked... The thought crossed my mind that I should look at it. Then [I thought], 'No, it's advertising cost information, not expense information. It probably won't give me any further information.' (3)

[Do you remember anything from the case on advertising?] "I remember underlining a few lines on it. It had something to do with the price of advertising. When I read it I thought, 'Price of advertising. I don't think that's going to be very important.' So I discounted it totally. I didn't even think about it." (42)

Unlike these two, another clicked on the button twice and interpreted what he saw as follows:

"Was advertising something? The button sits there just begging you to ask. It could have been one of the reasons for such wide variation." [Did you press it?] "Yes. And you [sic] saw nothing. It popped up a screen that said day/night. That was it. I didn't really pay much attention. And it disappeared." [What did it mean?] "Nothing. I interpreted it as this part of the game has not been implemented." (6)

Another, who understood the information content of the advertising window did not recognize that advertising was a variable:

[Did you think about advertising?] "I clicked on it once. This box came up saying that where advertised [sic]. Since I couldn't play with it, I just ignored it. (long pause) Oh, did it change over time? Once I couldn't adjust it, I (pause). Oh, that's true, it might have changed. It looked more like a setup to me, individual to me. [Do you remember what was in the box?] It was daytime versus primetime." (7)

These interpretations of advertising are idiosyncratic. The thread that runs them is not their content or the way they were formed, but the fact that they went unchallenged despite the low cost to participants of doing so.

Summary

This section has presented data suggesting that study participants adopted frames, both consciously and subconsciously, that affected the direction of their search efforts. Initially they drew inferences from the context of the task, specifically from the instructions and the case, that in some cases directed their attention away from the advertising variables. They limited the set of variables that they considered, thereby closing off the search for new
variables, in several ways. First, in an effort to control complexity, they would at times consciously restrict their thinking to the variables of which they were already aware. They also inadvertently focused on variables that they could control. They restricted their search for new variables to those that took a form similar to the outcome, a practice that made distinguishing the advertising variables less likely. Finally, they adopted idiosyncratic interpretations of the advertising variables' meaning. What is most remarkable about these frames is that sometimes they were formed with no data at all and they often went untested.

The incidence of each of these barriers among noticers is consistently lower than the incidence among non-noticers.

4.3.4. Discussion

The data presented above suggests three important elements that must be included in a theory of noticing.

The first is the motivation to search, which determines whether people initiate and continue search. The participants' behavior revealed three conditions other than intrinsic interest that helped determine the motivation to search. Their perceptions of whether results were possible from search played a role; if they judged that search would be unproductive, they gave up. Participants monitored different indicators of progress in the game; they focused on different aspects of the feedback stream (the size of the residuals, changes in the residuals, the nature of their distribution) or on the content of what they had learned (correctly or not) as variables that could be interpreted to indicate the need to learn more. More important than the measures themselves were participants' expectations about them. The fact that expectations sometimes drifted to match the status quo or were arbitrarily formed and left unchallenged suggests that the formation of expectations, rather than the measurement of feedback or progress, is the more important factor in determining whether or not people search.

Two factors emerge as critical enablers on the motivation side. The first, of course, is task motivation. People must be intrinsically interested in the task and have sufficient time and resources available to be able to engage in it. The second is an awareness of the arbitrariness of any absolute interpretation
of feedback, and therefore a skepticism about any conclusions drawn about the need for search and its likely rewards.

The second element of this partial theory consists of scripts for the search process. Often, participants' behavior implied scripts poorly adapted to this particular task. Participants who would generate hypotheses readily based on very few observations used the same script for hypothesis testing. These tests, based on few observations, led to binding conclusions that acted as constraints on participants' further thinking. Their scripts for handling "confirmed" hypotheses did not appear to respond to the strength of the evidence supporting the hypothesis.

Furthermore, there were errors in the way participants handled disconfirmed hypotheses. Occasionally they would respond to disconfirmation of a hypothesis by abandoning the associated predictor variable entirely. They confused the validity of variables with the validity of the hypotheses of which they were a part. A large number of the participants exhibited scripts that constrained severely the range within which they would experiment, leading to a conservatism that made it difficult or impossible for them to learn.

A few participants resorted to an "I've missed something" script when they became frustrated. These participants reread the instructions or the case to try to find information that they had missed or incorrect assumptions they had made. In many cases, this type of reconsideration led subsequently to noticing. Given that participants generally appeared to exhibit a bias toward working with the variables with which they were familiar rather than search for new ones, this script was an enabler.

The third element consists of frames, which acted as limits on the range of variables considered (i.e., the content or focus of search). The data suggest that participants adopted frames that constrained the problem, drawing inferences from the instructions and sometimes arbitrarily limiting their view to their current sets of known variables. Participants also excluded from consideration variables that they could not control or, based on untested assumptions about the meaning of the advertising variables, excluded them as well. Participants sometimes looked at but failed to distinguish the
advertising variables, perhaps because they relied on a similarity heuristic that was unlikely to associate the numeric outcome with the textual/categorical advertising-related predictors. Participants rarely appear to have questioned the assumptions underlying each of these ways of constraining the search for new variables. In many cases, they do not appear to have been aware of them.

The data suggest a process of search which potentially leads to noticing that is similar in form to an anchoring-and-adjustment process. It starts with the adoption of some context- or experience-supplied model. Search then proceeds to explore relationships between variables that are known as candidates. Only if this search is frustrated (or if no variables are provided initially) does the process go to the third step, the search for new variables. A final step, apparently rarely achieved, is the examination of the scripts and frames guiding search that themselves may function as limitations on the process.

Search as a hierarchical process

Overview

The data suggest that model construction proceeds as a kind of anchoring with three types of adjustment. The process starts with the adoption of some context- or experience-supplied model whose most important characteristic is a specification of the candidate set. Search then proceeds in three directions. In order of preference, they are:

1) Search within the candidate set to construct models and test them
2) Search for new variables to add, one at a time, to the candidate set
3) Search to identify scripts and frames that constrain or bias the other two search modes

These three types of search are hierarchical. People resort to the second only after the first has failed; and they return to the first as soon as they notice a new variable. The third level is reached only after the first two have failed.

The initial model

People infer from the situation they face and, by analogy with their past experiences, an initial model. Initial models can also be supplied by other
people, as well. In the experiment, the linear price-quantity relationship was strongly implied by the case as well as by the participants' own experiences with introductory economics. People who enter the process with no initial model are forced to begin by searching for an initial, candidate variable.

The initialization process described here results in two outcomes, an initial model and, potentially, frames that subsequently guide the search process. The formation of both is marked by what Einhorn and Hogarth (1982) term the fluency of our reasoning. They write:

"A significant feature of causal/diagnostic thinking is the remarkable speed and fluency which people seem to have for generating explanations and accommodating discrepant facts into expanded hypotheses... Since diagnostic thinking is so fluent, one must guard against the way cues to causality quickly restrict our interpretation of the past." (Einhorn & Hogarth, 1982: 32)

Participants in the study, seeking to orient themselves, were highly susceptible to the contents of the instructions, even where they were imprecise. The language in the case regarding advertising is purposefully vague; yet it appeared to influence a considerable number of people in the study. In fact, people are likely to use any source of information and in so doing were unlikely to weigh what they learned by the reliability of the source (Tversky & Kahneman, 1982), a phenomenon analogous to their subsequent failure to weigh the strength of the data supporting their conclusions from tests of hypotheses. In a conversation about the difficulty of constructing and controlling the context in which experiments take place, John Sterman recalled discovering that people playing "The Beer Game" behaved differently depending on the size of the pile of chips (used to represent cases of beer) left on the table; were it not for this psychological link, the size of the pile would have been irrelevant. Within organizations, people are told explicitly, through job descriptions and procedures, and tacitly, through their interactions with others, that certain information, actions, and goals simply are not their responsibility. One person's incorrect and offhand remark that something is irrelevant or relevant, may result in a permanent constraint on another's thinking; the "don't worry" exhortation in the case is something we all have heard, perhaps gratefully, and most likely failed to question. Thus initialization processes can be powerful in organizational settings.
Furthermore, where these frames may be supported by an official organizational logic, or are the product of habit and informal "truce" (per Nelson & Winter, 1982) between people about what are and are not their responsibilities, they are likely to be more or less permanent. In organizations, then, frames are likely to be supported by political structures and departmental boundaries, reinforcing mechanisms not found in the lab, and thus may be even more difficult to question. The directions in the case, then, are simply a metaphor for the melange of signals we receive from our environments. Our unhinging receptivity to them and their rigid subsequent use are phenomena noted in the lab and in others' research (Chanowitz & Langer, 1981; Snyder & Swann, 1978).

*Search within the candidate set*

Despite the fact that in most cases determining the variables that make up a system contributes more to performance than determining the particular form of the functional relationships between them (Dawes & Corrigan, 1974; Klayman, 1988), people prefer to work out the relationships between the variables that they know about.

Brehmer (1980) discusses a hierarchy of relationships that people tend to hypothesize when considering associations between variables. It starts with positive linear, then negative linear, and eventually ends in relationships with no pattern that must be memorized and finally probabilistic relationships. The participants in this experiment often resorted to memorized relationships that were complex, arbitrary, and, on their face, bizarre and unlikely. Almost all acknowledge the role of probability in their models, invoking "noise" or "error" and attributing time dependencies to it or relating it to other variables (typically quantity). Whether explaining the role of variables such as price as predictors or specifying the role of error, participants called on arbitrary and therefore memorized relationships to help explain their observations.

It is quite clear that people have powerful tools, both for constructing models from known variables and for interpreting feedback creatively, that allow them to make sense, or feel that they are making sense, of situations without expanding the set of variables they consider.
Search for new variables

Generally, given enough feedback that they interpret as disconfirming, people will search for new variables. When they have no initial model, of course, they must engage in this type of search. Given an initial model or just a list of candidate variables, however, they will not begin by searching for new variables. They must first be confronted with what they perceive as a failure of the known variables to predict.

Brehmer's (1980) model does not extend beyond exploration of the relationships between known variables. Based on the results from the experiment, I propose that people search for single variables to help them predict the unexplained variance they experience. People anchor off their initial model and adjust one variable at a time. This mode of adjustment, that is, one variable at a time, makes it very unlikely that people will identify interactions between variables; since the search for new variables focuses on single variables, those without strong predictive validity (main effects) of their own are unlikely to enter the candidate set. Given the difficulty people have learning to use interactions (Camerer, 1981; Sterman, 1994), I hypothesize that if the search for single variables fails to identify useful predictors people may be more likely to give up than to initiate a search for multiple variables whose interaction serves as a predictor. What was termed above a preference for unidimensional solutions may represent a preference for single-variable explanations of variance over ones that involve multiple variables or interactions.

Search for constraining scripts and frames

Only occasionally will people examine their own search behavior to identify the constraints that they have created for themselves in the form of scripts and frames guiding search. Even when participants in the experiment were aware of these constraints, as was several times the case when they realized the limitations associated with how they were recording data, they often chose not to change their behavior. Thus awareness was only the first step. They needed to go further than identifying the frames by recognizing the constraints they implied and then deciding whether to continue searching within those constraints.
Summary

Based on observation of the participants in the experiment, I have constructed a model of the inference process and identified a set of phenomena that influence it, either enabling or preventing noticing.

The model states that people build models through a hierarchical process. They start with hypothesis generation and testing using variables that the modeler already thinks are relevant (the candidate set). Only if a satisfactory model cannot be built from variables in the candidate set is a search for new variables invoked. If this search, too, fails, then search can escalate to a search for constraints acting on the search process itself: an attempt to identify constraining assumptions in the form of scripts and frames. As soon as a higher-level search yields results (i.e., a new variable is added to the candidate set or some invalid assumption is discovered), search returns to the next lowest level; thus only one discovery (new variable or constraining assumption) is likely to be identified at a time.

The barriers and enablers fall into three categories. The motivational barriers and enablers determine if the process takes place at all. Intrinsic motivation and an awareness of the necessary ambiguity in many problem framings play an important role. Script-based barriers and enablers affect activity within each type of search and oversee the process. At a low level, scripts affect hypothesis generation, hypothesis testing, and the removal of variables from the candidate set. At a higher level, they determine when and if people engage in the two higher levels of search in the model. An important high-level enabler specifies searching for potentially invalid assumptions when frustrated. In a complex world where one's senses are overloaded, frames might act as important enablers and contribute to noticing. Here, however, they acted primarily as barriers. Frames focus attention on certain types of variables and thus play a role primarily in directing the search for new variables.

Noticing of new variables depends on these barriers and enablers. It requires adequate motivation to search – a product of intrinsic motivation and of frames that view the problem as both unsolved and worth trying to solve. Search scripts must be used that exhibit a lower threshold before the
search for new variables is invoked. In the face of ambiguous feedback, people may need to invoke search scripts for finding new variables even when their instincts direct them to think within the current candidate set. Finally, even when the search for new variables occurs, we need to guard against the constraints of our own framings, searching not only for variables, but also for the frames we may be using inappropriately to direct our search processes.

4.4. Barriers/Enablers and the Explication-Noticing Relationship

Thus far, this chapter has been primarily descriptive, identifying and illustrating potential barriers to and enablers of noticing. In so doing, it has painted a picture of the search process as guided by biased frames and constrained by frames that are often arbitrary, idiosyncratic, or based on flawed judgment. Worse, these frames once formed can be held unconsciously and quite rigidly, independent of the strength of the data on which they are based. I have proposed that these barriers and enablers mediate the relationship between explication and noticing.

What remains to be done is to demonstrate the proposed mediating role played by barriers and enablers. The analysis presented here has two components. The first creates a high-level, aggregate operationalization of the barriers/enablers construct by characterizing each participant on the basis of the number of barriers and enablers that he or she exhibited. It then uses logistic regression to test whether the barriers/enablers are associated with noticing and whether they mediate the explication-noticing relationship.

The second analysis is primarily qualitative. Rather than aggregating the barriers and enablers, it considers each one's relationship to explication. It also discusses the role of data-recording, demonstrating the very important contributions that it makes to performance that have not been acknowledged elsewhere in the analysis of the experiment.

4.4.1. The Mediating Role of Barriers and Enablers

In this section, I offer an analysis of the barriers and enablers identified in the section above. In it, I aggregate the barriers and enablers into a single variable – labeled "Barriers" in Tables 4.2, 4.3 and 4.4 – operationalized as the
aggregate number of barriers exhibited by each participant minus the number of enablers. The higher the number, in principle, the less likely the participant should be to notice. The distribution of this variable is shown in Table 4.2.

**TABLE 4.2**
Frequency Distribution of Barriers/Enablers

<table>
<thead>
<tr>
<th>Number of barriers exhibited</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>1</td>
<td>3 (7%)</td>
</tr>
<tr>
<td>2</td>
<td>13 (31%)</td>
</tr>
<tr>
<td>3</td>
<td>6 (14%)</td>
</tr>
<tr>
<td>4</td>
<td>8 (19%)</td>
</tr>
<tr>
<td>5</td>
<td>4 (10%)</td>
</tr>
<tr>
<td>6+</td>
<td>6 (14%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>42 (100%)</strong></td>
</tr>
</tbody>
</table>

The analysis below has been checked for sensitivity to different operationalizations of the aggregate barriers variable. Three operationalizations were used. The first is the variable as initially aggregated. The second eliminates the tail on the distribution by categorizing all participants who exhibited six or more barriers into the same group. The third does the same for all participants exhibiting three or more barriers. The different operationalizations do not alter the result of any analysis presented below.

Explicators exhibit fewer barriers to noticing than non-explicators ($F(1,40)=5.17, p<0.05$). The same cannot be said for data-recorders and non-data recorders ($F(1,40)=2.04, n.s.$). The means and standard deviations for these comparisons can be seen in the marginals of Table 4.3, below.

**TABLE 4.3**
Mean Number of Barriers, Summarized by Recording Behavior

<table>
<thead>
<tr>
<th></th>
<th>Explicated</th>
<th>Did not Explicate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recorded data</td>
<td>2.4 (2.3)</td>
<td>3.3 (1.4)</td>
<td>2.8 (1.8)</td>
</tr>
<tr>
<td>Did not record data</td>
<td>1.7 (0.6)</td>
<td>4.0 (2.4)</td>
<td>3.8 (2.3)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2.2 (1.9)</strong></td>
<td><strong>3.9 (2.2)</strong></td>
<td><strong>3.4 (2.2)</strong></td>
</tr>
</tbody>
</table>
The body of the table suggests that while explication has a main effect and data-recording may not, data-recording and explication might interact. A two-way analysis of variance including both main effect and the interaction, however, does not support this hypothesis. In it, data-recording and its interaction with explication are both non-statistically significant; their addition does not improve model fit. The best model includes only the main effect of explication.

The association between explication and barriers, however, does not demonstrate the barriers' mediating role. Models 1 and 2 in Table 4.4 show that explication and barriers are useful – and almost equally good – predictors of noticing. A comparison of Models 1 and 3 demonstrates that adding barriers to a model already including explication improves model fit ($\Delta$log likelihood = 7.20, df=1, p<0.01). Model 4 shows that data-recording is not a useful predictor.

| TABLE 4.4 | Logistic Regression Models Using Barriers and Explication to Predict Noticing |
|------------|----------------------------------|-------------------------------|-------------------------------|-------------------------------|
|            | Model 1  | Model 2  | Model 3  | Model 4  |
| Constant   | -2.16**  | 1.63~    | 1.15     | 1.21     |
| Subsample  | 1.65**   | 1.48~    | 1.89*    | 2.07*    |
| Explication| 3.25***  |          | 2.82**   | 2.63**   |
| Barriers   |          | -0.96*** | -0.74**  | -0.77**  |
| Data-recording |        |          |          | 0.87     |
| -2 Log Likelihood | 39.54  | 39.73  | 32.34 | 31.51 |
| df          | 2        | 2      | 3       | 4       |

*: p<0.05   **: p<0.01   ***: p<0.001

Although adding barriers as a predictor reduces the explanatory power of explication (the coefficient drops in magnitude), both remain useful predictors, as comparisons between Models 1 and 3 and Models 2 and 3 demonstrate. This suggests that barriers may mediate, but cannot fully account for, explication's effect. Therefore, the barriers and enablers identified in this chapter must represent only a subset of the mechanisms by which explication has its effect. There are still more remaining to be discovered.
4.4.2. Explication, Data-Recording, and Barriers/Enablers

Much useful information is lost in the analysis just presented because of the high degree of aggregation. In this subsection, I return to qualitative analysis in order to discuss more precisely the relationships between explication and the barriers/enablers. In it, I also consider the role of data-recording; I demonstrate contributions that data-recording makes to the process of model construction, specifically its affect on search within the candidate set, and propose that it may affect some barriers, as well.

Table 4.5 shows the barriers data completely disaggregated and summarized first by explication and then by data-recording. This table serves as the basis for the discussion that follows regarding the associations between recording behavior and barriers/enablers.

Motivation-Based Barriers

Generally, explication's effect on noticing is not mediated by motivation. Arguably, explication, because of its effect on performance, is part of the system of feedback described earlier that affects motivation. I propose that the strongest link, however, makes explication the combined product of the treatment and the participant's level of motivation; the link from explication back to motivation is far more indirect and diffuse. It is clear that as the motivation to search waned, so did explication; participants who gave up also stopped explicating. Thus, participants explicated while they were still motivated to search and the explication helped them notice; but explication's effect on motivation was at best indirect. The same can be said for data-recording; since it was spontaneous, it was almost purely a product of motivation.

Table 4.5 reveals no evidence of association between explication and either task motivation or giving up. This makes good sense. Task motivation was determined primarily by factors outside the laboratory and by the individual's response to the task. These factors would be uninfluenced by explication.

The fact that explicators were no less -- and perhaps even more -- likely than non-explicators to give up is due in part to the fact that explication was associated with noticing the advertising variables. People who noticed were
### Table 4.5
Frequency of Barriers, Summarized by Explication and by Data-Recording

<table>
<thead>
<tr>
<th>Type of Barrier</th>
<th>Explicated freq</th>
<th>Explicated %</th>
<th>Did not explicate freq</th>
<th>Did not explicate %</th>
<th>Recorded data freq</th>
<th>Recorded data %</th>
<th>Did not rec. data freq</th>
<th>Did not rec. data %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motivation-Based Barriers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Low task motivation</td>
<td>1</td>
<td>9</td>
<td>3</td>
<td>10</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>• Giving up</td>
<td>4</td>
<td>36</td>
<td>7</td>
<td>23</td>
<td>3</td>
<td>19</td>
<td>8</td>
<td>31</td>
</tr>
<tr>
<td>• No perceived need to improve</td>
<td>1</td>
<td>9</td>
<td>8</td>
<td>26</td>
<td>4</td>
<td>25</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td><strong>Any motivational barrier</strong></td>
<td>4</td>
<td>36</td>
<td>14</td>
<td>45</td>
<td>6</td>
<td>38</td>
<td>12</td>
<td>46</td>
</tr>
<tr>
<td><strong>Script-Based Barriers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Problems interpreting small numbers of observations</td>
<td>1</td>
<td>9</td>
<td>12</td>
<td>39</td>
<td>3</td>
<td>19</td>
<td>10</td>
<td>38</td>
</tr>
<tr>
<td>• Drops vars when hyp. disconfirmed</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>10</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>• Restricted range of experimentation</td>
<td>1</td>
<td>9</td>
<td>13</td>
<td>42</td>
<td>3</td>
<td>19</td>
<td>11</td>
<td>42</td>
</tr>
<tr>
<td><strong>Any script-based barrier</strong></td>
<td>1</td>
<td>9</td>
<td>20</td>
<td>65</td>
<td>5</td>
<td>31</td>
<td>16</td>
<td>62</td>
</tr>
<tr>
<td><strong>Frame-Based Enabler</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Re-examination of own frames</td>
<td>3</td>
<td>27</td>
<td>3</td>
<td>10</td>
<td>4</td>
<td>25</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td><strong>Frame-Based Barriers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Frames from the context</td>
<td>3</td>
<td>27</td>
<td>14</td>
<td>45</td>
<td>7</td>
<td>44</td>
<td>10</td>
<td>38</td>
</tr>
<tr>
<td>• Fixing the set of candidate variables</td>
<td>5</td>
<td>45</td>
<td>12</td>
<td>39</td>
<td>6</td>
<td>38</td>
<td>11</td>
<td>42</td>
</tr>
<tr>
<td>• Excluding what you can't control</td>
<td>1</td>
<td>9</td>
<td>5</td>
<td>16</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>• Frames constraining form of data</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>45</td>
<td>4</td>
<td>25</td>
<td>10</td>
<td>38</td>
</tr>
<tr>
<td>• Specific (untested) interp. of adv.</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td><strong>Any frame-based barrier</strong></td>
<td>7</td>
<td>64</td>
<td>26</td>
<td>84</td>
<td>13</td>
<td>81</td>
<td>20</td>
<td>77</td>
</tr>
</tbody>
</table>

Faced with complexity and participants with complex models of the situation were more likely to give up. In this case, giving up did not mediate the explication-noticing relationship but was an outcome of it.

---

10 Since the totals reflect number of participants and some participants may exhibit multiple barriers, the "any barrier" line is not necessarily the sum of the disaggregated lines above.

11 Since the category "looking at advertising only once or twice" is not confound with either explication or data-recording it can safely be included in this Table. It is confounded with noticing and therefore was excluded from Table 4.1. This also explains why the total number of participants in the any-script based barrier category differs between the two tables.
The table does reveal a negative association between explication and low perceived need to improve. In general, explicators were less likely than non-explicators to become complacent about their current levels of performance; they did not "get used" to the status quo. It is possible that the treatment was a strong signal to participants that they were expected to learn something and thereby improve their performance; however, the interviews suggest that virtually all participants expected to learn during the exercise. It is also possible that explication changed the way participants formed their performance expectations, thereby preventing acclimation to the status quo. Finally, explication may have helped raise participants' awareness of their own lack of mastery, driving them either to give up or to continue search. It should be noted that although explicators may have been more likely to expect performance to improve, their expectations were no less arbitrary or idiosyncratic than non-explicators'.

My conclusion is that explication has a role in the feedback system involving motivation described earlier in this chapter (see the summary to Section 4.3.1). While it is unlikely to have any influence on task motivation, through its effects on framing it may affect giving up and the perceived need to search. The data suggest that it has its strongest effect on the latter. The loop is clearly closed, however, since as the motivation to search waned, so did explication; participants who gave up also tended to stop expediting. Because of the many causal links here and because these links sometimes act in opposite directions, I argue that motivation does not play an important role in mediating the explication-noticing relationship, at least in comparison to the script- and frame-based effects.

With respect to data-recording, my interpretation of Table 4.5 is that motivation tends to drive data-recording, particularly since it was a spontaneous response among participants. It may be associated with (not) giving up, however, I suspect that both are a product of task motivation. Finally, the table offers no evidence of an association between data-recording and the perceived need to improve. My conclusion about data-recording is that it is primarily a product of, rather than an influence on, motivation-based barriers and enablers.
Script-Based Barriers

As the script-based barriers portion of Table 4.5 shows, there is an association between explication and each of the script-based barriers discussed in this chapter. As compared to explicators, non-explicators were more likely to exhibit problems interpreting small numbers of observations, to drop valid variables entirely when invalid hypotheses were disconfirmed, and to restrict tightly the range in which they experimented. Overall, explicators were almost 7 times less likely than non-explicators (9% versus 65%) to exhibit script-based barriers to noticing.

The bulk of the people exhibiting these barriers were non-explicating, non-data-recorders, which raises the possibility that data-recording might be responsible for the observations. A comparison on the basis of data-recording reveals that data-recorders are about half as likely to suffer from any one of these barriers than those who did not record data.

It seems likely, of course, that both explication and data-recording would have an effect. But they may operate quite differently. I would argue that explicators are just as likely to draw conclusions from small numbers of observations as are non-explicators; however, I propose that they are more aware of the judgments they make and more likely to question and reconsider them. Explicators, for example, often offered incorrect hypotheses about the roles that variables played and did so on the basis of essentially no data. However, their diaries and interviews are full of instances when they corrected these tentative assertions. Non-explicators were more likely to cling to tentative results and to have them dominate their thinking throughout the experiment. All three participants who dropped variables entirely because a hypothesis was disconfirmed were non-explicators. When participants made their thinking explicit, they also were able to challenge it. They were perhaps no less likely to fall into conceptual traps but they were much more likely to become aware of them and escape.

Data-recording most likely worked differently. Rather than helping participants to become aware of and to challenge their assumptions, it alters the thought process by tying participants more closely to data. It makes them less dependent on their potentially biased memories and perceptions, and
perhaps makes them more aware of the strength of data supporting their hypotheses.

I expected that data-recording would be very important for avoiding conservatism in the range of experimentation. It is not clear that it had this effect or that it was any more effective than explication in reducing this type of conservatism. It appears, however, that only a very few participants actually tried to calculate the appropriate size of the price adjustments they should make; most, even those who recorded data, used trial and error, focusing on the direction (either bigger or smaller) much more than on the magnitude of their adjustments.

**Frame-Based Barriers**

The likelihood of occurrence of almost all the frame-based barriers to noticing seems lower for explicators than for non-explicators. Furthermore, unlike the script-based barriers, there is generally no association between them and data-recording. Each of these represents a conceptual constraint on the direction of search. Gleaned from the context, from past experience, or from casual inference, these frames represent untested, often arbitrary assumptions about the kind of information likely to prove useful.

Although explication appears to have had no effect on participants' likelihood of fixing the set of candidate variables, it reduces substantially the likelihood of all the other frame-based barriers. Explicators were also particularly likely (compared to non-explicators) to challenge assumptions they made about the advertising data based either on experience or on their initial interpretation of the case. Furthermore, explicators simply did not fall into the trap of looking at advertising and dropping it because they failed to see numbers in the window; they were more flexible in terms of the form of the information to which they would respond. Finally, explicators (including the two non-noticing explicators) did not report any untested assertions about the meaning of advertising and therefore its irrelevance.

When examining the incidence of any frame-based barrier for explicators versus non-explicators (64% vs. 84%), it is important to recognize that fixing the causal set often occurred after noticing advertising. If it is left out of the summary, the difference between explicators and non-explicators is more
pronounced (27% vs. 81%)\textsuperscript{12}. Explication, then, appears to have a powerful effect on frame-based barriers. By contrast, these barriers seem unaffected by data-recording.

4.4.3. Discussion

Explication and the Noticing Process

I propose that explication acts in two fundamental fashions. The first is that it encourages, in fact requires, the formation of new frames or hypotheses. Second, by making these models explicit, the explicator also makes them accessible so that he or she may challenge them. Thus participants are less likely to use them unthinkingly as rigid constraints on search. It seems that the explicators had many of the thoughts that appear to have constrained non-explicators, but that these thoughts did not constrain them in the same way. They were more likely to question their ideas.

However, explication did not address all the barriers identified in this chapter. The data suggest that explication had little, if any effect on task motivation. Its motivational role, if any, is to decrease the likelihood of being satisfied with current performance levels.

Explication appears to lower the likelihood of occurrence of all the script-based barriers; its effect, however, is confounded with that of data-recording, making it difficult to attribute explanatory power to either variable. Overall, explicators were only one-seventh as likely as non-explicators to exhibit a script-based barrier to noticing. Explication appears to play its most dramatic role in reducing the incidence of frame-based barriers to noticing, making it far less likely that search will be constrained by arbitrarily formed and untested frames.

Although not shown in the tables, it is worth noting that nine participants had realizations during the interview about variables they had not considered in the simulation exercise. They mentioned them spontaneously and were surprised by their behavior. This suggests that the explication taking place in

\textsuperscript{12} The frequencies are: explicators, 3 (27%); non-explicators, 25 (81%). For data-recording, 10 (63%); non-data-recording, 18 (69%).
the interview may have had an effect similar to the explication taking place on paper. In 7 of the 9 cases, the participants remembered the instructions but questioned why they had not looked at advertising anyway.

The quantitative analysis presented in Section 4.4.1 helps demonstrate the role that the barriers and enablers play in mediating the relationship between explication and noticing. The analysis also demonstrates that they cannot fully explain explication's effect. There are more mechanisms remaining to be discovered. Specifically, the analysis has identified relatively few enablers and I think that this is the most promising area for future inquiry. Explication might lead to greater creativity, clarity of thought, use of analogy, creation and consideration of alternate scenarios, or counterfactual reasoning, all potentially important enablers that did not emerge in this study.

The Role of Data-Recording

The role of data-recording seems quite different from explication's. Data-recording makes past data available for drawing inferences, thereby reducing memory-related biases that can affect judgment. It is likely to reduce superstitious learning and it lowers substantially the effort required to retest a hypothesis that gets called into question. In short, data-recording should increase the quality of judgments made in testing hypotheses.

However, data-recording may also serve as a constraint, focusing attention and thus hypothesis generation on the data that have been recorded. In any situation where data-recording is institutionalized or routinized, it may have this constraining effect. In the experiment, data-recorders adjusted what they recorded on the basis of their current hypotheses and sometimes were constrained later by those decisions; in a sense, their data-recording was customized to the task of testing a particular hypothesis and generally became less useful if and when their hypothesis was disconfirmed. Participants who explicated in addition to recording data did not appear to constrain their thinking on the basis of the data they had recorded.

Data-recording seems to play a critical role in interpreting feedback accurately and parameterizing a model once it is formed. If we examine score rather than noticing as the outcome, and tabulate it by noticing and data-recording (as in Table 4.6), we can see the influence of data-recording. Given
that explication is associated with noticing it comes as no surprise that it has a main effect on score. However, data-recording has a dramatic effect on score among noticers. Without it participants parameterize their models only poorly; with it, they do much better\(^\text{13}\).

<table>
<thead>
<tr>
<th>Score, Summarized by Noticing and Data-Recording</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Recorded data</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Noticed adv.</td>
</tr>
<tr>
<td>Didn't notice</td>
</tr>
</tbody>
</table>

I propose then, that data recording's primary effect is in parameterizing models, that is, in search within one's candidate set. It helps most once one has formed a model that includes the right variables. It plays a secondary role in model formation by helping increase the accuracy of hypothesis testing, reducing somewhat both Type I and Type II errors in testing the relationships between variables.

Alternate Causal Paths

I have presented a single interpretation of these data. Yet I believe that there are at least two other explanations and a second causal path that makes participants' reports a consequence of whether they noticed. In this section, I review the two unaddressed interpretations.

The first argues that some of the frames participants describe in interviews are in fact post hoc constructions, rationalizations used to justify behavior rather than the actual causes of those behaviors. The second suggests that participants who notice and those who do not are likely to remember different things and thus report events quite differently. I offer evidence for both phenomena below and argue that they do not fully explain the results. They leave much room for my primary hypothesis, that frames did direct participants' search efforts and that the framing process occurred differently among explicators and non-explicators.

\(^\text{13}\) Adding data-recording and its interaction with noticing to a model of score that includes only noticing results improves model fit (F(2,37)=4.45, p<0.05).
Justifying one's behavior

Nisbett and Wilson (1977) argue that we have poor access to our own mental processes. Thus, when asked to justify our attitudes or actions, we cannot simply recall their causes. Instead, we infer them, at times offering explanations that are justifiable in some social calculus but have little to do with their true causes. If the conclusions of this research are true, it is quite possible, if not likely, that some of the retrospective descriptions offered by participants were post hoc justifications of behavior that they might not even understand.

"I was getting frustrated... so I reread the instructions. I read through the things that whoever said to look at. I realized I never looked at advertising..." (She looks in her notes) "Yeah, there it is, week 27." [Did you use it?] "Well, it was pretty general (pause) I did use it... it didn't seem to work." [How long did you use it for?] "Probably 10 weeks." [Why weren't you using it until week 27?] "Probably because it said it wasn't relevant or not to worry about costs. So maybe in my mind I thought advertising costs. And then I realized, 'Why's it there? It must be there for a reason.'" (8)

The phrasing of the last question is unfortunate because it is likely to create a demand effect calling for justification. Prior to this question, this participant had discussed the difficulties he had had trying to use the advertising variables. Not once did the case or instructions emerge as important. Nor did they emerge as part of the reason to stop looking at advertising. In such a case, it seems likely that references to the instructions were a justification rather than a cause of behavior.

Another participant offered a series of explanations for not using advertising during the course of the interview:

"[In discussing the screen: Did you click on this button?] "Nope." [Any reason?] I didn't want to ruin the game. If I'd been there by myself or if it was a class assignment, I probably would have." [It would ruin the game?] I don't know. This was a structured game. I'm doing something for you guys. If it had been a class I would have clicked on everything." [Did you worry you were going to break it?] "No. I just

---

14 I maintained notes on interview questions that I felt were leading or would otherwise bias participants. Mistakes such as the use of this question were not repeated.
didn't know what was going on. It's your thesis. I didn't want to mess with someone else's thesis." [Then later] "The case did say not to worry about advertising. So if it said not to worry about advertising yet, that's another reason I didn't touch the advertising button." (5)

My impression of participants who offered multiple explanations for not looking at advertising was that they were generating explanations for their behavior rather than reporting ideas generated during the simulation.

One participant who never looked at advertising reported:

[Did you think about the advertising?] "No. When I saw that the slightest changes in price were generating such wide variations in demand, I didn't think it would have an effect. I just didn't think about it. Did it change from period to period?" (11)

It seems unlikely that the person who asks if advertising changed from period to period in the game would also have thought carefully about its relationship to variation in demand. My feeling is that the variation argument is rationalization and that this person "just didn't think about it."

Contrast these last people's explanations to the following two:

[Did you think about advertising?] "I wasn't supposed to mess with advertising, was I? I wasn't paying any attention to it?" (15)

[Did you click on the advertising button?] "No, I didn't. (long pause) That's funny. You said in the instructions you don't even need to click on it." (17)

The first was surprised by my question about advertising. The second was surprised at his own behavior; he remarked later that it was "...amazing that I never hit the advertising button to check it out."

There were participants who were clear that advertising was irrelevant to the problem but could not remember why they thought so. Others offered explanations that were so logically unsatisfactory that it is hard to imagine constructing them as justifications:

"...I tried it once but it didn't work out." (34)
It seems unlikely that the MBA with a few statistics courses under his or her belt would construct such a statement to justify anything. Could a report that failed so miserably at face-saving actually be a justification? Or was it a spontaneous report about a casual (and flawed) judgment process? My instincts suggest it is the latter.

Participants often made no attempt to justify their inferences; they just stated them. Their surprise and spontaneity in discussing advertising suggests that they were in fact guided by their casual interpretations of the case.

At times, however, participants' explanations appear to have been justifications. In discussions of specific variables other than advertising, they would often say that they had not even thought of them. Typically, people who claimed to have thought about but not used advertising almost always tried to explain why. In some cases it was clear that they tried to use the advertising variables, failed, and moved on to consider other variables or decision rules. In others, however, I could find no evidence that participants had tried to test the hypotheses offered; their explanations were inconsistent with their behavior, suggesting after-the-fact justification.

It is clear that, at times, participants did construct justifications for their behavior. However, while justification may have been responsible for some of participants' reports in the interviews, the phenomena described in this chapter are real and not just the product of justification.

**Schema-consistent recall**

The schema-consistent recall hypothesis (see for example, Cohen, 1981; Lord & Foti, 1986; Ross *et al*, 1975) would suggest that noticers and non-noticers, having framed the problem differently by the end of the game, would be likely to remember different things when interviewed; their memory would be biased toward ideas or memories that are consistent with their schemas. Noticers, therefore, would be less likely to remember the references to advertising from the case (which suggest that it might not be useful) since they are inconsistent with their mental models of the problem. Further, among people who remembered references to advertising, the
noticers would be less likely to remember that the directions seemed to imply that advertising might not be useful.

Participants who remembered a specific reference to advertising indicating it was not useful said things like:

"She said, 'Don't worry about advertising.' I didn't know if it was a trick for me not to look too soon." (20)

"For some reason I wasn't supposed to mess with [the advertising]. I think that's what the woman in the case indicated." (15)

Others remembered only that advertising had been mentioned but no more:

"I looked back for it [advertising] and I saw in the instructions that it was a variable." (14)

"I'm trying to remember where I got this information about advertising [that it was a variable]. Where did I find it? I looked back but I couldn't find it. It said something about advertising but I don't know what." (35)

"It didn't mention anything about advertising in the case. In the instructions it just said that it was there." (39)

Table 4.7 shows that, overall, noticers were less likely than non-noticers to remember references to advertising from the case. Only 8 of the 16 noticers (50%) remembered a reference to advertising versus 19 of the 26 non-noticers (73%). This difference, however, is not statistically significant ($\chi^2_{(1)}=2.30$, n.s.). Among the 27 participants who remembered any reference from the case,

| Table 4.7 Remembering References to Advertising in the Case and Instructions |
|---------------------------------|---------------------------------|----------------|---------|-------|
|                                 | Remembered adv from case/instructions | Did not remember | Total  |
|                                 | Don't worry / not useful | Neutral reference |                |
| Noticed advertising             | 3 (19%) | 5 (31%) | 8 (50%) | 16    |
| Did not notice                  | 17 (65%) | 2 (8%) | 7 (27%) | 26    |
| Total                           | 20 (48%) | 7 (17%) | 15 (36%) | 42    |
non-noticers were more likely to remember that the case indicated advertising might not be useful. 17 of the 19 non-noticers (89%) had such memories compared to only 3 of the 8 noticers (Fisher's exact test p=0.01139).

These data, particularly those that relate to the content of what was remembered, are consistent with the schema-consistent recall hypotheses. On the other hand, they are also consistent with the idea that the people who were most influenced by the case were also the least likely to notice. If all the data collected were retrospective, then it would be difficult to distinguish between these hypotheses. However, it is clear from notes that participants in the treatment group took and from other participants' decisions and data-recording behavior, that participants were influenced by the directions and, for that matter, that each of the limiting frames and judgment errors presented earlier in this chapter did in fact occur.

Summary

We must acknowledge that the forces of self-justification and schema-consistent recall are real, are likely to affect self reports from participants, and therefore may have played a role in the data reported here. However, the questions in the interviews were carefully sequenced to avoid creating demand effects. Participants own responses suggested genuine surprise at their behavior; often they were at a loss to explain it. Participants' statements were checked against process data recorded during the simulation and against their protocols if they were in the treatment group. In short, the data supporting this chapter's analyses were carefully collected and tested at each step in the analysis. The conclusions are well supported.

4.5. Summary and Conclusions

In this chapter, I have proposed a partial theory of the noticing process. The theory consists of a model of the inference process combined with three sets of barriers and enablers that influence it and therefore make noticing more or less likely. The first set of barriers and enablers relate to the motivation to search, the second to scripts that help choose search activities, and the third to frames specifying the types of variables likely to be considered. Motivation determines if search actually takes place, scripts direct the process, and frames exercise control over content.
The analysis demonstrates that these barriers and enablers mediate the relationship between explication and noticing. It appears that explication's effect is due first to a reduction in frame-based barriers, then to script-based enablers and perhaps to a reduction in script-based barriers. Motivation-based barriers and enablers seems to be the least important group for explaining explication's influence on noticing, although they clearly are important.

The analysis also shows that the known barriers and enablers represent only a subset of the mechanisms through which explication has its effect; there are more barriers and enablers to be discovered. I proposed several enablers that, while they did not emerge from the data, might be worth examining in future research. These include creativity, clarity of thought, the use of analogy, the use of alternate scenarios, and the use of counterfactual reasoning.

The analysis demonstrates that data-recording aids people's accuracy in parameterizing the models they form but that it does not appear to play a very important role in the noticing process. I hypothesize that it could, however, by making hypothesis testing more accurate and encouraging inference on the basis of past data. There is a risk to data-recording. Data-recording can lead to institutionalized systems for recording data which ultimately function as frames restricting search and therefore noticing (see Sterman, 1994, and his discussion of NASA and measuring ozone levels for a vivid example).

The associational nature of the data raises the question of the direction of causality. Self-justification and schema-consistent recall both might have played a role in the results by providing a causal path from noticing to the results obtained. However, given the process-tracing from the experiment, the types of questions asked in the interview and their sequencing, and the varying forms these phenomena took in different participants (thereby creating an opportunity to triangulate), it seems reasonable and safe to conclude that explication operates through the mechanisms described here.

Before closing the chapter, I wish to emphasis three additional points. Each is an area that should be investigated in the future. The first relates to the use of small numbers in judgment. While researchers who study judgment and decision-making have documented the errors that people
make on the basis of the small numbers bias (e.g., Tversky & Kahneman, 1982), they neglect the critical role it plays in hypothesis generation. To succeed at a task such as this, participants often must generate many hypotheses; reasoning based on very small numbers is the engine for this process. However, participants who do not explicate tend to adopt hypotheses prematurely and then restrict their search efforts in accordance with them. Explication appear to help people use the hypotheses they have formed yet not be constrained by them in the same way. This phenomenon needs to be investigated in greater detail.

Second, this section has identified a judgment error that has not been studied: discarding variables when hypotheses including them are disconfirmed. Although the hypotheses may be incorrect, no inference can be drawn that the independent variables in them are not predictors. We need to examine more closely the circumstances under which people "un-notice" variables, that is, discard them from their sets of candidate variables.

Finally, we also need to examine the circumstances under which people add complexity to their mental models (e.g., interactions, non-linear relationships, appeals to random error or time-dependencies) rather than begin searching for additional variables that might play a causal role. The participants in this experiment were willing to construct and test complex, idiosyncratic models involving wild non-linearities, time-dependent error terms, and even relationships that were learned rather than functional (e.g., if A then B, if C then D, etc.). They were often convinced they were right and, despite the complexity of their theories, failed to wonder, "This is awfully complex; are there any other variables I should be looking at?" The model of the inference process developed here is governed by a script that determines when people search within their candidate sets and when they seek to expand them. We need to understand better how people choose between these two paths and then examine whether explication influences that choice.
5. Explication and Search Outside the Laboratory

5.1. Introduction

This chapter reports the results of field work to identify variables that were controlled by the experimental task and the laboratory setting but play roles in explication and noticing in real-world settings. It is important to recognize that this study was designed before the analysis of the experimental results was completed. As a result, the design was not influenced by the framework presented in Chapter 4. However, I use that framework to analyze the search-related behavior observed in the field and to criticize the intervention itself. After updating the framework to address new barriers emergent in the field, I conclude by offering a different approach to the intervention in response to the updated framework.

5.1.1. The Nature of the Experiment

In order to understand the goals of the field study, it makes sense to review first the conditions under which the results reported thus far were obtained. Each feature of the situation raises questions about the circumstances under which explication might occur and might be effective in producing noticing.

Characteristics of the Task

The task was a quantitative problem based on a relatively simple model where the predictors were powerful relative to error variance. Many of the problems that people encounter in their working lives are less quantitative, more complex, and provide predictors less powerful than those used here. Does explication become an even more powerful aid to noticing when models are complex and less quantitative or where individual predictors have explanatory power that is small compared to the error in the system? Might it function in the same manner as the "cognitive" (or "policy") feedback used in studies of multiple cue probability learning (Hammond et al, 1973)?
Chapter 5

The simulated environment contained no feedback loops, no nonlinearities and no interactions between variables. Furthermore, it had no time delays built into it. People have great difficulty learning about systems that embody these characteristics (Sterman, 1994; Diehl & Sterman, 1993; Paich & Sterman, 1993). Might explication help people identify and respond to these structural "variables" more readily than they would otherwise?

From the participants' perspective the experimental task was new and generally unfamiliar. Although they brought substantial relevant knowledge to it, they still were functioning in a new environment. Tyre and Orlikowski (1994) suggest that learning in such situations continues for a period of time but then stops as the situation becomes familiar and behavior routine. Future learning then occurs only when some event opens up a new "window of opportunity" for learning which, like previous ones, also closes after a period of time. The experiment captured participants in an initial window of opportunity; I would also argue that for many, we saw that initial window close. This raises several questions: Does explication help only when one is learning a task initially? Does it help during any window of opportunity for learning? Can it help create a window? Can explication have any effect on people working in environments with which they are intimately familiar and which, from their viewpoint, have been stable for years? Can it help people who have overlearned their jobs to notice new variables that have recently taken on some as-yet-unrecognized importance?

The experimental task had a fixed length; it was not an ongoing task. Furthermore, its duration was relatively brief; few participants worked at it for more than an hour. But most people's jobs, or at least their relationships to their work environments, are ongoing. Can explication be used on an ongoing basis? Is it a one-shot deal, or perhaps a tool that one can use occasionally?

The form of the Data

We know from the experiment that participants appear to have trouble distinguishing text variables in a task where all others were numeric. But variables can take on many forms other than those present in the experiment. They may take forms as diverse as actions people take, reports printed on paper, and sounds. They may all be observable at the same time and the same

- 160 -
place or may require one to remember past events or search in different places. Can explication help people learn to respond to variables that manifest themselves in unfamiliar ways?

**Characteristics of the Treatment**

The treatment was a written protocol that periodically asked participants highly structured questions (e.g., scale items about confidence), some structured but open-ended questions (e.g., "What variables are you thinking about?") and some completely open-ended questions (e.g., "Is anything else important?"). But it is actually a very complex intervention and therefore may have multiple and complex effects on study participants. I hypothesize that the treatment's most important characteristics are that it calls for people to state the variables in their mental models (i.e., those that should relate to the price they set) and that it encourages them to speculate rather than to report only inferences held with very high confidence. Were the request different, for example if it called for people to record only conclusions that they felt were correct, would it yield different results?

One important characteristic of the treatment is that it calls for writing on people's parts, not speaking. We know from the debate on the validity of verbal protocols that protocols are likely to be reactive unless administered under very specific guidelines. In contrast to the wealth of literature on the role of verbal protocols (Ericsson & Simon, 1980, 1984; Nisbett & Wilson, 1977; Payne et al., 1978; Wilson et al., 1989), there is precious little written on the topic of written protocols. In this case, explication proved quite reactive. But might verbal protocols be as reactive if used in the right ways? We know that participants in the interviews had spontaneous realizations, noticing the advertising variables and others, as well. Were the interviews an example of the potential power of verbal explication? Could a treatment calling for verbal explication be as effective as the intervention used in the experiment?

The topic of verbal protocols raises an additional issue. In general, verbalization is used for communicating with others and thus is a social act; writing or note-taking can be private. I discuss later the social issues that may arise. Here I merely wish to state that one consequence of asking people to verbalize, even when they are alone, may be to activate schemas with social content that might not be invoked when they write. Even if they know that
their verbalization will be confidential and private, people generating verbal protocols may experience imaginary but schema-induced social pressures.

Characteristics of the Sample

The sample consisted of Sloan Master's students. I discuss specifically the match between their skills and the task below. Here, I stress other characteristics. They were well educated and accomplished people, highly motivated and confident. They expected to solve the problem. In the words of one participant, "I wanted to excel." As I have pointed out earlier, motivation around the task and around explication was generally high. Would explication be effective with a less motivated group or one with lower expectations for themselves?

In part because the task was new and because their participation was temporary, participants had little if any stake in making decisions in any particular way. While they entered into the task with much potentially applicable knowledge, they were not personally identified with the task or the procedures surrounding it.

Skills and their Match with the Task

One obvious area to consider is skills. The participants in this experiment had taken microeconomics at the college and graduate levels. They all had taken at least a single semester course in probability and statistics. They all had at least one full year of an MBA curriculum under their belts; none of the variables presented to them should have been conceptually challenging. The task required participants to interact with a computer; however, they were all experienced computer users. In short, the participants had great familiarity with the general domain of the task and had skills appropriate to addressing it.

In situations where people's skills do not match the task so well, learning would likely be slowed. Would explication help those who are less familiar with the conceptual material of the task? Would it encourage the distinguishing of new variables or their movement from the candidate set to the known set (i.e., the process of linking variables to mental models). Would it encourage both?
Chapter 5

The (Experimental) Setting

Experimental settings have their own demand effects. Here, one important effect was the participants' near-ubiquitous assumption, at least initially, that there was something for them to learn, some way for them to improve their performance. This situation, like many real-world situations, provided no absolute reference for people to evaluate whether they were doing well or poorly, whether there was something they had not noticed, or whether there was more to learn. But the demand effect essentially eliminated this ambiguity; although I have noted exceptions, most people who failed to notice the advertising variables were convinced that they were missing something.

Such demand effects are not necessarily present in everyday work situations. There are two reasons (at least) why this might be. The first, familiarity with the situation, has been discussed above. The second is framing as communicated by others. Even offhand comments from others can produce important framing effects. Examples of this are many participants' responses to their fictional boss' comments about advertising in the case. Can explication help challenge frames created in this manner? Even when they are maintained socially or institutionally?

Another characteristic of experimental settings is that if you ask study participants to do something, in general they try to comply. Participants in the experiment did not appear to question why they should maintain written protocols; they just did what they felt was requested. Such compliance may not be easily obtained outside the lab. In addition, given that the amount of explication dropped notably in the last quarter of the game, it might be difficult to obtain compliance over longer periods of time or in situations where the need for search is less clear. It seems likely that there are many variables, controlled in the lab and therefore very difficult to notice, that represent barriers to the implementation of any explication-inducing intervention in organizational settings.

The Meaning of the Treatment in a Social Context

Up to this point, I have regarded the treatment as a request to engage in a behavior (that I call explication) as participants go about some task. I have not
examined further the way in which content or the source of the request might affect its meaning to individuals. Yet my comparatively neutral request to speculate and list variables is likely to carry different implications than the same request worded differently, for example as a request for participants to justify their decisions. The request was framed as coming from me, the experimenter; yet the instructions could have framed the request as coming from the participants' fictitious boss. The source of the request and the intended use of anything the participant produces may both affect dimensions of meaning that the laboratory study could not reveal.

Consider, too, the timing of the request. In the experiment it was made initially, as part of the instructions and as part of the task. What might people think if such a request were made part way through? They might interpret it as an indictment of their judgment rather than a neutral request, even with identical wording. At the very least, they would view it as something added to the original task rather than a natural part. What are the effects of the timing of explication and the degree to which it is perceived as integral to the task?

**Summary**

While the experiment offers convincing evidence for a causal relationship between explication and noticing, speculation about the uniqueness of the laboratory setting, the sample, and the specific task suggests variables that may either affect explication itself or moderate the relationship between explication and noticing. Furthermore, the ease with which potentially relevant variables come to mind raises the additional question of how many important variables remain unnoticed.

**5.1.2. Goals of the Field Study**

The study described in this chapter is an effort to identify variables forming the context of the relationship between explication and noticing. This context includes both variables that moderate the relationship between explication and noticing and ones that affect explication itself and therefore help determine the effectiveness of any intervention.

Such a search has both theoretical and practical relevance. For the theorist, simply demonstrating the link between explication and noticing
only begs the question, "Under what circumstances?" For the practitioner, the results from the experiment prompt inquiry into actions that might be taken to produce explication as well as the conditions under which explication would have its effect on noticing.

Although constructed to have some face validity, the experiment belongs in the realm of the theoretical. This study leans toward the practical. Rather than produce a complete and integrated theory of explication and its relationship to noticing, it seeks to explore the practical issues that managers might face in an attempt to use the phenomenon seen in the lab. It represents an effort to notice the variables that might form such an integrated theory.

The study has two central research questions:

- What are the variables that managers might have to respond to or adjust in order to produce explication?
- Were they to do so, what variables might interfere with explication's effect on noticing?

Secondary questions include:

- Given these variables, what might be an effective explication-inducing intervention?
- How might the nature of the circumstances affect one's choice of intervention?

This study, then, is primarily exploratory. Any consideration of results and or costs can come only after a successful implementation. Given the surprises revealed by the experimental piloting process, it seemed likely that initial attempts in the field might produce more learning for the experimenter than the participants, and might be short on results in terms of noticing. Hence the emphasis on exploration over experimentation.

5.1.3. Chapter Overview

In the next section (5.2), I describe the methods used in the field work. Section 5.3 describes the setting, focusing first on my observations about noticing and explication and then describing the study participants' response to the intervention. In Section 5.4, I analyze the response to the intervention in two ways. First, I examine what the intervention meant to the
participants, developing the notion of barriers to explication, a topic that never emerged in analysis of the experiment. Then I analyze the intervention's effect by calling on the barriers framework developed in Chapter 4. In this analysis, I identify issues emergent in the field and add them to the framework where appropriate. The chapter concludes with a summary of the learning from the field study and a proposal for another intervention that addresses better the barriers to noticing observed at the site.

5.2. Methods

The genesis of this study is a conversation I had with one of my touchstones in the practical world of day-to-day business, a production manager at a manufacturing firm. I briefly presented the experiment's design and results then asked him if he, as a manager, felt that this result was important. His response, cited on the second page of Chapter 2, suggests that he did. I then asked how he would respond to what he had just learned and if these results would affect his behavior in any way. "What, if anything, will you do now that you know this?" I asked. His response:

"Well, you know me. I'm an empirical kind of guy. I'd probably want to try it out."

And so a project was born. From Ben's\(^1\) point of view, my inquiry into the variables affecting the implementation and effectiveness of explication-inducing interventions was less important than the positive results that, in his opinion, might come out of this work. But he realized the size of the leap from lab to shop floor and recognized the risks it implied at least for short-term positive results.

5.2.1. Design

This research is an exploratory field study using a group of operators of highly automated, robotic assembly equipment. It was executed in two parts, an initial observation period and a second period marked by an explication-inducing intervention.

---

\(^1\) Each participant in the study was assigned a unique pseudonym to maintain strict confidentiality. Gender was not necessarily maintained in assigning the names.
Chapter 5

My goals for the initial observation period were to become familiar to the organization being studied, to develop a rich picture of the context in which I would be intervening and thereby notice potentially relevant variables, to explore the degree to which noticing on the operators' part seemed possible, and to learn how much explication already took place as part of the way this organization worked. I also intended to work with the production manager and supervisor to design the intervention. My goal was to employ their help in adapting the laboratory intervention and instructions to create an analog appropriate for their organization.

In the second period, I implemented the intervention and collected data on the results. In the period after the intervention, my principal goal was to observe how the operators responded to the intervention in an effort to identify variables affecting explication and its relationship to noticing.

I considered the idea of a quasi-experiment, using randomly assigned control and treatment groups. However, the sample was small (22 operators), participation by no means guaranteed, and the factors that might interfere with the effectiveness of the intervention unknown. Therefore, I decided to apply the treatment to all subjects and use a before/after design. This design virtually guaranteed relevant learning on my part while a quasi-experiment based on my then-limited knowledge of the contextual issues surrounding explication-inducing interventions in organizations posed considerable down-side risk.

5.2.2. Intervention

The intervention was developed with the production manager during the initial observation period. Ben's primary role was as a sounding board for my ideas as I discovered issues during the initial observation period. While he supported the project and met with me frequently, he did not contribute as much to the intervention as I would have liked. He did, however, help during implementation, again acting as a sounding board on implementation issues and offering advice.

The intervention took the form of a verbal request for help that I made to operators on a one-to-one basis. I asked if they would at some point each day take some time to think about problems in the room that affected
performance or product quality, to speculate on their causes, and to take notes on their ideas in a notebook that I would provide. The causes could be technical – issues with the robots, with their programming, the way they were designed, or maybe issues with parts. They might be people problems (I told them I didn't care about names, just the issues) or organizational issues. I told them they didn't have to be sure about their ideas; they could report pure hunches.

Nor did the diaries have to be pretty or even very legible because I planned to review them with the person once each week. I told the operators the purpose was to get them to take a little time out for thinking and to help make sure they remembered their ideas. For people who had difficulty writing in English, they could use their own language if they preferred, since we would go over what they wrote together. I told them that the important thing for me was that they spent some quiet time thinking, that they speculate, and that their ideas were recorded well enough that they would remember them.

In terms of time, I was not asking for a lot. I was hoping each person would get a chance to think about it and take some notes for perhaps 5 or 10 minutes a few times each week. I told them that if they could do more, that would be great and I would appreciate it. I also told them that their supervisor was aware of what I was doing and that he authorized them to spend as much time at this as they wanted (all of which was true). The initial plan was to do this for four weeks and see where we were at that point.

I told each person that participation was voluntary, that I would not reveal to anyone the names of the people who participated, and that I would not reveal the sources of any statements that I quoted. If at any point in the process they were uncomfortable with what they were doing, they could stop their participation.

I did not offer a reward or pay for participation. In the initial observation period, operators were very open. Anything that I could offer seemed trivial and perhaps even belittling. Instead, I appealed to them for help and proposed that through it they might learn. At the very least, it would prove interesting.
Chapter 5

I asked if they understood what I was asking them to do, if it made sense to them and if they wanted to help out. I gave each person who volunteered (a total of 12 of the 22 operators) a notebook with a one page summary (see Appendix E) of what I had just told them. I also suggested that when they actually tried to do it, they might realize that my instructions were not so good; if they had questions or wanted some help they should ask me. In saying that I tried to put the "blame" for difficulty doing the task on me rather than them.

5.2.3. Data Collection

Data collection took place over a seven-week period. I was around the factory approximately three full days each week. Of my time spent there, the vast bulk was spent on the lines, observing and talking informally with operators or engineers as they went about their work. I attended meetings, virtually all of which were weekly and production related, although there were a few that were specially called.

The primary tools of data collection were:

- My own observations as recorded in a field notebook and daily transcribed and further embellished into a set of field notes
- Informal interviews with all study participants
- Interviews with engineers about their interactions with the operators.
- The notebooks that volunteers maintained (i.e., their written explication)

I also examined all documentation that operators maintained. These documents consisted of production logs and several types of production data used for analysis.

5.3. Results

In this section, I describe the setting and the ways in which people responded to the intervention. The description of the setting has three components. First, I describe the operation, the people who worked in it, and their roles. The next part discusses operators' opportunities for noticing and establishes that noticing was in fact possible. The third part uses descriptions of the operators' daily life to indicate the degree to which operators explicated
naturally in the course of their work. The section closes with a description of how operators responded behaviorally to the intervention. These responses are analyzed in depth in Section 5.4.

5.3.1. The Setting

The setting was a production assembly area in which machines assembled intricate electromechanical devices, similar in scale, complexity, and parts mix to a miniature tape recorder (e.g., a dictaphone or a walkman). The production tasks were highly automated, the vast bulk performed by four robotic assembly lines, each consisting of between 7 and 24 robots assembling a separate subassembly of the final product. The assembly operation was quite ambitious from a technical standpoint because the parts, which included very small springs, tiny screws, and flexible plastic circuitry, and snap-fit parts, were all very difficult for robots to handle and put together. As a result, the lines were highly sensitive to adjustment problems, variation in the dimensions (and sometimes even the finish) of the parts, and the presence of dirt.

The lines occupied a single large room populated by 22 machine operators supported by 4 engineers, a supervisor, and my principal contact, the production manager. The entire factory, including the robotic assembly area, was supported by several electricians and mechanics. None of these people were unionized and all had worked in their roles for at least a year, most for longer.

Generally speaking, operators were familiar with all the tasks within the operator role on a given line. They switched roles through daily or weekly rotation. The tasks were similar from line to line and many operators had worked on several. Operators carried several responsibilities. First, they handled trays of parts, placing full ones into staging areas from which the line could draw a supply as needed and removing empty ones that had passed through the system. Second, they removed completed subassemblies and inspected them. Some of the lines automatically inspected every subassembly; operators would remove those that were rejected, diagnose the problem, and repair them if possible. Finally, the operators were responsible for watching the line, catching problems, perhaps diagnosing them, and
reporting them to an engineer, mechanic, or electrician. Very occasionally, operators would fix problems. There were two with greater technical expertise than the others. They moved from line to line helping to diagnose and sometimes to solve problems. Despite these exceptions, operators were called upon primarily to execute procedure and secondarily to identify and report problems.

Unlike the operators, who constantly attended to the line and to immediate production needs, the engineers focused on solving problems. Since their time spent on the line was devoted to specific problems, they relied on the operators or on reports to help them identify problems. They also were proactive, making improvements to the line to improve quality, reliability or speed. They also helped manage problems with incoming parts. The mechanics and electricians ensured that the equipment was maintained properly and they fixed it when it broke; but unlike the engineers they did not make changes to it.

5.3.2. Noticing in the Observation Period

I had three primary questions about noticing to be answered in the initial observation period: whether the environment provided variables to be noticed, whether the operators were in a position to do that noticing, and whether the operators actually did any noticing.

Did the environment contain variables to be noticed? Unlike the laboratory, where I knew what variables were causal and thus what valid variables there were to be noticed, in the field I did not. On reviewing the performance history of the lines, I realized that the performance of the two largest of the four lines varied considerably from day to day and typically performed well below what should have been possible from a technical standpoint. The engineers responsible for the lines assured me that there certainly were many outstanding problems that had not yet been identified. They claimed that whenever they worked on the line, they would discover more. The most obvious symptoms, robot stops, were familiar. It was in the process of diagnosis where discoveries might be made. They noted that among the problems that had been identified, there were always several on each lines that had not been diagnosed; the engineers were so busy they could
not keep up with their work. They fixed the most important problems within a matter of days; but less dramatic problems could remain unaddressed. Engineers agreed, then, that at any time there was an inventory of performance or quality related problems that had not been discovered, and another group that were known but had not been diagnosed. Recognition and diagnosis of these problems rarely offered up new variables for the engineers who, in general, knew the robots inside and out; however, the operators, whose knowledge of the way the robots worked was less detailed, stood to learn a great deal. From the engineer's viewpoint, the operators had substantial opportunities to notice new variables.

These problems were caused by several types of variables. Some were due to equipment design, including the design of the control software. Occasionally a problem would arise whose cause was attributed to product design. Variation in the incoming parts could also create assembly problems on the line or for quality. Then there were problems that arose because people failed to perform their jobs properly or because some activity had had an unanticipated side-effect. Finally there were situations where the subassemblies produced on the line were returned from other departments but where the cause actually lay in some part of the process downstream of the four lines; the causes of such problems had an organizational element to them. Typically, problems were the product of multiple domains: too much variance in the parts combined with a process design that was not robust to the variance. Were the parts more consistent, or the process designed a little differently, or the equipment maintained and calibrated a little more carefully, the problem might not have occurred. The point is that diagnosing problems and designing solutions often required thinking in several different domains of variables.

Were the operators in a position to do that noticing? There are two dimensions to this question, one related to skills, one to opportunity. Engineers and operators both agreed that operators' knowledge of these different variable domains was limited to some extent, but that they were capable of a great deal of diagnosis. Many of the variables that emerged during diagnosis were not ones that required an engineering background to understand. Furthermore, the variables that needed to be noticed often were
not root causes; they were symptoms useful for diagnosing the problem, often in the form of detailed descriptions of behavior (e.g., "I think that the robot is picking up gears when they are in the tray upside down. I think it used to skip upside down ones before."). Such phenomena, once identified, could be represented unidimensionally (e.g., present vs. not present); nonetheless, they seem to provide a good contrast to the comparatively simple and obviously unidimensional variables used in the experiment. Importantly, noticing these behaviors was more likely to require a watchful eye rather than any special technical skill. If engineers pointed these things out to operators and asked them to watch for them, the operators had no difficulty doing so. Measurement was usually not difficult and the variables themselves did not require operators to think in fundamentally new ways. In short, operators had the skills to observe many of these variables. Furthermore, the operators, by virtue of their near-continuous involvement in day-to-day operations, were particularly well placed to notice these variables.

It is important to remember that noticing includes two fundamental operations, distinguishing new variables and linking them into one's model. Engineers reported that while there were opportunities to distinguish new variables, these events were relatively rare. I was aware of it taking place only once during the study. What was more commonly needed, in fact on a daily if not continual basis, was the second part of noticing: linking known variables into the candidate set for some particular problem. Forming these links often could be difficult because of the conceptual, physical, or temporal distance between cause and effect (all cues to causality per Einhorn & Hogarth, 1982). Causes might be removed physically, perhaps located at a vendor site. They might be conceptually distant as was the case for parts that were damaged in the receiving and material handling process. Sometimes they were removed in time as was the case with maintenance operations that were skipped and led to problems weeks later. In this environment, an extensive known set had been established; and even the operators possessed much of it. Because the operators had limited knowledge, noticing did call on them to distinguish new variables (certainly more so than it did of the engineers). However, it would be fair to say that linking rather than distinguishing dominated the noticing task.
Chapter 5

Were the operators noticing? The short answer is: not much. But operator noticing was critical to problem solving. Earlier work (Rabkin, 1992) reported that engineers at this site relied heavily on operators to identify and report problems. The pattern revealed in the earlier work was still present in the observation period. The engineers spoke both of their dependence on the operators and their need for more diagnostic information. They wanted operators to notice and report problems earlier, before they became substantial enough to make waves on the weekly production reports, before the number of stops they produced rose to very high levels. This, of course, required noticing symptoms beyond the obvious. The operators were in a position to notice these variables. Yet, the engineers indicated high levels of dissatisfaction with the amount of information they received from them.

Of course the measure of noticing implied by the previous paragraph confounds operator noticing with operator reporting. Perhaps the operators were noticing but simply were not reporting, or acting in any way on, their findings. I spent extensive periods of time with operators. I would often spend two hours continually with an operator while he or she worked. They would tell me what they were doing, what was bothering them, why somebody else was getting the first break, how to do the things that they did, what was wrong with the company, that they weren't paid enough, why they weren't paid enough, etc. Given the number of comments that were prefaced by "between you and me..." or "I'm not mentioning any names but by now you probably know who I'm talking about...", they did not appear to be censoring their thoughts. In short, most operators were open. I asked enough questions about the process and things that I saw going wrong to which they did attend that they knew I was interested in how things worked (or didn't work), performance problems and quality issues.

But, with the exception of two operators, I observed very little noticing. It was rare to observe operators initiating any systematic diagnosis (e.g., examining a machine every time it stopped rather than just clearing the part it failed to place and restarting it). On only three occasions in the observation period did an operator say to me the equivalent of "Now, why did that happen?" or imply such thinking through the actions he or she took. In short, it is my opinion (and one shared by the engineers, the supervisor, the
manager, and several of the operators) that there was a general deficit of noticing and search on the part of most operators. In the next section, I tackle the question of why.

I think, then, that it is fair to say that this production environment did provide opportunities for operators to notice. Even though the bulk of the work was routine and the vast majority of variables familiar members of an established candidate set, both engineers and operators claimed that completely new variables emerged periodically as important. More importantly, variables that were distant to the problems, typically upstream in the production process or actions earlier in time, had to be linked into the models of particular problems. Distinguishing variables was infrequent (i.e., weekly or even monthly) compared to linking variables into models (that is, recognizing known variables as members of the candidate set for a particular problem), which was called for almost continually.

5.3.3. Explication in the Observation Period

In the experiment, I could be assured that participants in the control group were not explicating. But in the field, I anticipated that operator's daily lives might provide opportunities for explication. If so, I wanted to identify this behavior and see how common it was. In what follows, I use incidents recorded in my field notes to put together a composite picture of the range of operators' activities. As we shall see, the operators explicated little.

The shift starts at seven o'clock. I was down at one end of the room talking with Bernie as he and Michelle started up Line 1. The other operators on this line, seven in total including the line's "lead" operator, had not arrived yet. They drifted in over the next 20 minutes or so. No one was discussing any issues with the line. I asked Bernie about this. Did they get together at the start of the day to discuss what would happen?

"When we had three shifts, there was a time when we would have meetings at the beginning of every day. The C shift supervisor and a few of the operators would get together with us – all the A shift operators, our supervisor, Ben, the engineers, and their manager – and we'd talk about the problems that had come up overnight... After a while, the managers wouldn't always come, Ben wouldn't be able to show up sometimes. We kind of let the whole thing slide."
was the last time you had one of these meetings?] "I don't know. At least a year."

He discussed those meetings as a time when pretty much everybody would speak up and where they focused on issues. For a time, they kept a flip chart with messages that helped communicate between shifts. However, in the time I was there, the flip charts were not used except for occasional notes about who was taking what break or instructions from engineers about what batch of parts to use first.

At this point Ben and Earl were at another meeting, where production quantities, quality, and issues of coordinating the different parts of the plant were discussed. But there were no meetings for the operators, no gathering to lay out the strategy for the day, surface problems, or make special requests.

Instead, as the morning wore on, engineers would pass through the room, making requests or offering warnings. One morning Alec came in and announced that the final assembly operation (a manual assembly area outside of and downstream from the robotic assembly area) had identified a quality problem linked to a bad part. He asked the operators to remove all assemblies with the bad part, stockpile them while he figured out what should be done with them, and to start running with a new batch of the part. Such a request was unusual. More typical might be an engineer's request for the operators to make sure they finished off a batch of parts before introducing any from a new batch and to inform engineering when they switched. Communication at the start of the day was haphazard, occurred only on an ad hoc basis, and, with very few exceptions, took the form of directives from engineers to operators.

There were two types of regularly scheduled meetings whose content and participants overlapped somewhat. Once each week there was an "operations" meeting attended by Ben, Earl, a variety of engineers, people from other departments in the organization on an as-needed basis, and an operator from each line. This meeting reviewed the performance of all the lines in the room. Three weeks into the study, my field notes read:

I asked Earl if he was headed to the operations meeting. He responded, "I don't know if there's going to be a meeting today since Ben's out.
This is Ben's meeting. But let's go over to [the conference room] and see if anyone shows up." ...another typical operations meeting. No one showed from Line 1 or Line 3. Alice and Ann showed from their lines... Ben was out.

The two operators sat together at the far end of the conference table separated from everyone else (except me, who also sat in the back) by several seats. They were silent as the engineers reviewed the week's production data. The engineers and the supervisor briefly reviewed slides showing production quantities and what the top quality issues were. The only effort to elicit participation from the operators came at the end of the meeting after the engineers had wrapped up. "Do you ladies have anything to say," Earl asked. They did not.

Operators participated minimally in these meetings. I never observed anything even remotely resembling explication on the operators' part in these meetings. Attendance was irregular; I never attended one of these meetings where all four lines were represented. The attendance problems were not restricted to operators, however; these meetings took a back seat to other things that were going on for virtually everyone involved.

There were also weekly meetings devoted to each line. Unlike the operations meetings, which may have been attended sporadically but did take place each week, the line-specific meetings often did not take place at all. However, one week I attended the meeting for Line 3. It was attended by Earl, several engineers, and two operators. Irma, an operator, had a special assignment and had been tracking the quality problems. She brought to the meeting a slide listing the top quality problems for the past week and the frequency associated with each. She put the slide on the projector but neither led the discussion nor presented the slide. Earl and the engineers discussed the problems.

This specific meeting is notable because it included the only instance of spontaneous participation by an operator that I observed in one of these meetings. The top quality issue was a chronic problem, present for more than a year, that was the result of procedural problems downstream in the manufacturing process. The end result was that roughly 20% of Line 3's output was being sent back as defective from a downstream operation that
was, in fact, responsible for breaking them. For various idiosyncratic reasons, the department that was breaking the parts had incentives not to solve the problem or to respond to appeals from Line 3. Ed, the engineering manager, announced that the time had come for him to go talk with the plant manager about this problem. At this point Ellen, the second operator in the meeting, asked, "This has been going on for more than a year and we've complained before. [The plant manager] is leaving so what does he care? (Although it had not been announced, everyone knew that he was moving on to another job in another month or so.) They didn't do anything about it before, why will this be different?"

Later in the meeting I found out why Ellen was there. At one point in the meeting, the engineering manager turned to her and asked, "So what do you have on dirt." She had been asked to do some testing to see how dirty certain piece parts were and had a slide for each of four parts. Basically, she would take a batch of the parts and shake the dirt from it onto a copier and then copy the result onto a slide. The slides had little black dots on them. In less than two minutes, we had run through Ellen's slides with no analysis on her part. The issue of dirt was not a problem Ellen had discovered or about which she was concerned. She had been asked to generate the slides and had been doing so weekly for several months. When I asked her whether there were issues with dirt and how she had gotten responsibility for generating these slides, she said that she had been asked to generate the slides by the engineers. As far as she knew it was the engineers who had raised concerns about dirt; it certainly wasn't she.

While the regular meetings may have provided an opportunity for operators to raise problems and speculate about their causes, the fact is that few operators attended or were asked to do so. When they did, they participated when specifically asked to do so and only rarely as a result of their own initiative. Even then, there is no evidence of any explication on their part.

The weekly meetings discussed were the only regular meetings related to the functioning of the line. As I mentioned before, there were no meetings at the start or the end of the day, formal or otherwise.
During the course of the day and as problems came up, operators did maintain a production log. The logs documented any instance when the line or a robot had to be shut down. Each item would include the time the shutdown started and when it ended, a brief description of the problem, always on the order of ten words, and occasionally a description of what was done. A typical log entry might appear as follows: 10:25am-10:50am, Robot 4 stops repeatedly placing part X, sensor error, called electrician. The "sensor error" comment actually reflects the message that would come up on the robot's display and is far from a diagnosis. In fact, the logs were virtually devoid of diagnoses. Furthermore, they did not involve any speculation; they documented what had transpired. If they did contain a diagnosis, it was entered after it had been verified and corrected. Thus production logs were used to document what happened in retrospect. Even then, they were incomplete because the operators often did not learn from whoever fixed it what had caused the problem or how it was fixed.

Operators also recorded information relating to product quality. Those who repaired failed subassemblies keep sheets with each type of defect listed and put a check mark by each when it was encountered. Those who inspected subassemblies that came off the line would sample them randomly, inspect the sample, and record the nature of any failures. The information that was recorded was standardized, a series of categories and frequencies.

When specifically asked by engineers, operators might record data needed to analyze problems. I observed this behavior on five occasions. In all cases, the data recorded was the same, a routine matter. In one case, the operator initiated it; in the others, engineers initiated it by asking operators to keep track of the data.

There were no other formalized opportunities for explication. There was no meeting at the end of the shift. However, there were some informal opportunities that arose in the course of each day. These arose typically through interaction between engineers and operators and were often triggered by a chance encounter. For example, one such incident occurred when one of the operators had noticed a potential problem, a part that was defective (it was a plastic part that was not complete because the mold was not being filled completely with molten plastic). She had not raised any alarms or
Chapter 5

asked anyone about it. But when Alec happened to walk by, she caught his attention and showed him the part. She asked him if the defect was a problem. He told her no, but to keep an eye on it. If it got worse they might have to switch to another batch of parts. In this case the operator did not explicate at all. However, sometimes operators did do some diagnosis.

At one point, Ellen called in an engineer, told him that a given stop had happened a series of times. She had not tried to diagnose it and offered no possible cause. However, she saved one of the parts that the robot could not place and gave it to him because "who knows, something might be wrong with it." Later, referring to the part she had saved and the information she gave him on the incidence of the stop, the engineer told me, "Even that is a lot more [information] than I usually get."

These chance interactions with engineers provided opportunities for operators to report their ideas about what the problems were and what might be causing them. Although there were a few notable exceptions, operators' interactions with each other and with engineers did not include explication. Operators rarely tried to diagnose problems and when they did, they would report a single theory about what was happening without stating their assumptions or offering any alternatives. They reported only what they believed was correct; they did not report tentative hypotheses, a critical element of explication. We shall see this same pattern emerge among the few operators who attempted to explicate in response to the intervention.

In short, this organization provided no formal structures that supported explication. Informally, explication was supported but rarely realized through day-to-day interactions between engineers and operators. The observations made in the initial weeks of the study reveal very little explication and the pattern continued throughout.

5.3.4. Outcomes of the Intervention

Of the 22 operators in the group, 10 immediately agreed to participate in the study by taking notes. Two wanted to think about it and subsequently agreed to participate. Of the remaining 10, only 2 actually said no. The others wanted to think about it and avoided making a decision. I approached each one three times over the period of a full week, each time indicating that their
participation was voluntary and their decision confidential. Eventually I stopped asking.

Of the twelve operators who ultimately agreed to participate, only three actually took any notes in their notebook. A fourth claimed that she did not have time to write but had done some thinking. With her, and with others who approached me wanting to talk with me about problems on which they had not taken notes, I asked them to spend a little time writing before we spoke. In the final week of the study, I relaxed this policy simply because I was curious to learn what they were thinking. On talking with this fourth person, it was clear that she had in fact been speculating on causes of poor performance and had been attempting to think quite broadly.

The other eight participants offered excuses for why they had not done what was requested. I will discuss these at length in the subsequent section.

The intervention apparently resulted in dramatic changes in my relationship with only one operator, one of the two who flat out refused to help. Although he had been quite willing to talk with me prior to my request, after it he became belligerent. While prior to the intervention, he would greet me when I arrived, after he would say sarcastically, "Not you again" or "What do you want this time." One time when I was speaking with another operator, he passed by and said, "Don't forget to put a commendation in each of our files." I reminded him of the discussion we had had on confidentiality. I will discuss some of the issues that emerge from his response to the request and how they fit into patterns exhibited by other operators. The other operator who openly declined to participate had always been reluctant to speak with me. He would answer questions and make an effort to do so well; but he never initiated a conversation. With these few exceptions, my interactions with operators remained unchanged, even with those who avoided participation but would not say so explicitly. Those who had been talkative continued to be; those who had not been did not become less so, unwelcoming, or hostile.

As I mentioned, there was a very small group, three operators, who did take notes. These three wrote sketchy notes on a total of six problems, revealing a single cause for each. In no case did they offer multiple
Chapter 5

hypotheses about causality or reveal causes in which they expressed any doubt. They did not speculate in these notes or in our discussions about them. As far as they were concerned, they were right. There were two problems that had not been solved at the time we discussed them; for both, the operators expressed complete confidence that their diagnosis was correct. In one case the operator turned out to be correct. The other was never resolved (during the study); however, one engineer expressed doubt that the operator was correct and offered several alternative (and in his opinion more likely) explanations.

I believe that the operators engaged in a version of explication so limited that it should not be termed explication at all. Documentation might be a more appropriate label. They did express mental models of the causes of phenomena that they observed. However, they offered no speculation and expressed no uncertainty. They reported beliefs, not hypotheses, and did not express any uncertainty. They did not report their models as they were being constructed, but only after construction and testing.

This stark description of the results of the intervention fails to capture the factors that contributed to it. In the discussion that follows in the next section, I turn again to the barriers framework to understand the events that took place and organize the forces that shaped them.

5.4. Discussion

The prior section makes clear that while the environment provided stimuli to be noticed as variables, operators were not doing much noticing. I observed very little explication, suggesting that if an intervention could induce explication, it might make a difference. The intervention used, however, did not produce explication. In this discussion, I analyze the reasons why. The analysis focuses first on the different meanings that the intervention had for the operators and how those meanings may have affected their participation. Then I call on the barriers framework from Chapter 4 to analyze the barriers to noticing present in the setting and the intervention's effect on them. In the course of the analysis, several new barriers emerge that were not revealed in the experiment.
5.4.1. The Intervention and Its Two Levels of Meaning

The intervention took the form of a request made to operators to identify problems contributing to poor machine performance and decreased product quality, to speculate on what might be causing them, and to record notes on their ideas, no matter how speculative. In analyzing the events following the intervention, I discovered that operators interpreted the intervention to mean many different things. These interpretations fell into two categories, those that arose from the fact that I was making a request (and the form of that request) and a second that stemmed from what operators were being asked to do. This subsection develops these two levels of meaning.

The Meaning of a Request to Explicate

The particular request I made meant different things to the various operators and these meanings shaped their response. In this section, I discuss the implications of making a request of operators and the way in which that request was made. It does not address the meaning of the specific task that I asked them to do, a topic explored in the next section.

*The request called for operators to take initiative.*

Early on, I was concerned that if participation were made mandatory, operators would resent it. Thus I felt it was important that they be encouraged but not required by their supervisor to participate. Asking operators to volunteer, while it may have been sensible (and ethical) from the point of view of a researcher enrolling participants in a research project, was unfamiliar and unsettling territory for many operators. I heard the word volunteer used only once in the factory: an operator told me that she had "been volunteered" to do some data collecting by the supervisor. Operators were not given choices. They were told what to do.

Although operators complained about the situation they also took steps to ensure that they would not have to take initiative themselves. They avoided any work that was not well specified. If work had results that could not be measured unambiguously to indicate if they had or had not done it, they avoided it as well. Although there were clear exceptions, many operators did not take initiative to go beyond what was specified and required.
Management and engineering responded by specifying further the operators' jobs, thus closing a feedback loop. The work became more proceduralized. While this reduced the opportunity for operators to avoid tasks, it also continually demanded less in the way of initiative. What the operators had to do was specified and absolute; they did it or they weren't doing their job. The boundaries around their roles became increasingly rigid. The end result was a system that neither demanded nor rewarded initiative.

In the next subsection I will discuss the consequences of stepping outside one's role. Here, I mean to point out that operators weren't used to making decisions. My request posed a dilemma for many operators. It offered the opportunity to take some initiative. One operator offered up, "You know, you've got some nerve asking people to do something extra." When I reminded him that everything that operators did with me, or didn't, was voluntary and confidential, it only made him more annoyed. "I don't have time. I've got three jobs. I don't want to do it. Really, you've got some nerve." The fact that he may not have had the time, may have had three jobs, and may not have wanted to do it do not explain his anger. I propose that he was annoyed in part because I was asking that he make a decision and therefore take responsibility for the consequences.

Three operators wanted to talk with the supervisor before agreeing to anything. All three saw two possible outcomes. If he told them to do it, they would help me out. If he didn't want them to do it, they would not. In their conversations with me, none seemed to consider the third possible option, that the supervisor would be indifferent (which, in fact he was not; he wanted people to participate) and that he would tell them to decide for themselves. My feeling is that these operators were doing more than checking with the supervisor if it was OK; they wanted him to make the decision for them.

My attempt to not put pressure on operators, and thereby elicit genuine participation, turned out to be unsettling for many among a group of people unused to taking initiative and comfortable (although not necessarily happy) with being told what to do. By asking people to volunteer, I was also inadvertently and implicitly asking them to take responsibility for their actions.
Chapter 5

The request challenged role boundaries.

Operators who took my request were stepping beyond the boundaries of their roles, boundaries that I have already noted were carefully maintained. Although there were exceptions, operators typically viewed their role as doing what they were told to do by others, specifically their supervisor or, at times, engineers. Those operators who represented exceptions, that is, those who regularly took on their own initiatives, recognized that there were risks to stepping beyond their specific role.

Many operators referred to a lack of teamwork among operators and that they could not trust each other. One operator thought that another was "trying to set me up" because he (the speaker) had done something "extra that looked good." Another told me about an operator who stopped talking to her because she did some work that she thought was helpful but that the other operator felt was not part of the job; she said, "I'm a threat because I'll do a little bit more. They'll do only what they have to." Stepping beyond the boundaries of one's role created the possibility of conflict with others.

Without marching orders from their supervisor, agreeing to help me was an act attributable only to the individual. There would be no excuses, no way to deflect blame onto others, and thus no way to avoid any sanctions that sprang from participation. Without official notice stating that participating in my work was mandatory, and therefore part of the job, doing so was clearly beyond the scope of their work. Thus, when I asked operators to do anything unusual, I was asking them to step outside of their roles and take the risks associated with doing to.

The request raised uncertainty.

What may have been most unsettling for operators about my request was that the consequences of fulfilling it (or refusing to) were not clear. As I mentioned earlier, typically operators were told to do things. Occasionally they would object en masse. More likely, however, they would complain, comply to the extent that the request specified (and no further), and blame the consequences on whoever made the request. If they did not comply, they might lose their jobs. If they went beyond what was requested, they felt there
would be no special rewards and they would risk each others' wrath. Therefore, they did the minimum necessary to comply.

The request to explicate, however, came from an unfamiliar source, whose role and authority were not clear. I came and went as I pleased, spoke freely with anyone in the factory (in fact people would often stop by to say hi to me), and could attend any meeting that I wanted. Although I said in no uncertain terms that they were free to decide, that I would not tell their supervisor who did and did not help me out, and that there would be no consequences to them of their decision, the fact is that many did not trust me. Several operators wanted to know if I worked for the company (despite the fact that I had told everyone that I did not). Rumors about me would surface from time to time. "Is your father on the board of directors?" No. "Is it true that you are going to be replacing the plant manager?" No. Without understanding fully my role or the activity in which they were to participate, operators saw uncertainty in my request. Without knowing the consequences of saying yes or no, they were quite reasonably very uncomfortable in doing so.

Two asked me directly: "What happens if I don't help you?" My response was, "What happens if you don't have time to talk with me? Nothing. Our conversations, or lack of them, are confidential. So is this. I don't tell anyone." But most did not ask. Three operators asked what I would do with any notes that they took for me. I told them that I would review them with them and that I might copy them so I could look them over again but that I would never show them to anyone else. Near the end of the study, in explaining their response to the task, some participants told me that they did not know whether they could trust me: "We can't tell you everything. We don't know what you're going to do with [what we tell you]." When I reminded them of my promise of confidentiality, one said, "But you can't trust anyone." Another told me that they have been lied to many times and a third said that there was "...a lot of backstabbing, no cooperation, no trust."

To compound the problem, operators also felt that extra effort went unrewarded. They spoke of how the engineers did not listen to them, did not appreciate it when they did help out, did not thank them or commend them to management, and were not supportive when evaluating operators for raises and promotions. In the words of one operator (who I observed to be
very conscientious), "There's very little in the way of rewards for good behavior." This opinion was almost universally echoed by others. Thus there was no positive side to the uncertainty.

In the absence of trust, my request created uncertainty on people's parts and the uncertainty all had to do with risk. Most operators did not expect there to be any rewards.

Cooperation with the request singled people out.

Many participants, more than three quarters, wanted to know who else or at least how many others had agreed to participate. When I asked why they wanted to know, they responded that they were just curious. The first operator with whom I spoke said, "It will be interesting to see how people react to this." When I asked if she thought people would object, she offered a noncommittal, "I don't see why they would."

No operator wanted to be the only one participating. One, upon agreeing to help me out, immediately dragged me over to another operator, a friend of hers, and started telling the other person how they both were going to be doing this thing for me. She didn't give me a chance to talk or her friend an opportunity to decline. Perhaps the first operator was simply trying to be helpful to me. More likely, however, her actions represented an insurance policy against being the only person participating.

Operators did not want to be singled out. Those that were, who were asked to do anything out of the ordinary by their supervisor or an engineer, were scrutinized by other operators. Others would ask them or the supervisor why that person got special treatment; this happened even in situations where the task assigned was something the complainer had avoided in the past. Anything that differentiated one operator from another created potential problems for the operator who stood out, independent of whether it was assigned or the operator's own initiative. Cooperation with me was potentially such a differentiating variable.

The request represented an opportunity.

The company had a pay-for-learning program and it is clear now, in retrospect, that at least one operator hoped that participation was somehow
tied into the evaluation process. Near the end of the study, I was talking with
her and said, "Next week, I hope I get the chance to talk with you about," and
then paused. She interrupted, ending my sentence on a hopeful note, by
saying, "Certification?" My response was to remind her what I had said (and
written) earlier about confidentiality; but I realized that she had hoped that
somehow a raise or promotion would come out of all this.

Perhaps others felt this way. However, my guess is that most operators did
not see important rewards associated with me or my request. Most took me at
my word and those that did not saw risks rather than rewards in cooperating.

The Meaning of Explication

Up to this point, I have discussed the meaning of my request absent the
fact that I was asking operators to explicate. But the specific task had meaning
to the operators, as well, and it must be examined if one is to understand their
responses to it.

Earlier, I described the fact that formal opportunities for explication were
virtually nonexistenent. Few operators attended meetings and those who did
participated minimally. Written logs were restricted to recording data about
product quality or maintaining production logs that documented events but
involved no speculation or diagnosis. Informal opportunities existed but
were rarely used since operators generally did not attempt to diagnose
problems and did not pursue diagnosis jointly with engineers. In such a
context, then, what might the particular note-taking activity that I requested
mean to an operator?

Explication meant extra work.

Perhaps the most common response that I got from operators was that I
was asking them to do extra work and that they did not have the time. Many,
despite my efforts to convince them otherwise, saw little of intrinsic interest
in what I was asking them to do. To them, it was either extra work or a waste
of time.

Presumably there are requests that I might have made that they would
have viewed differently. One said, "You can come talk with me any time.
But I don't have the time to write things down." He was good for his word,
too; he was friendly and more than willing to talk with me while he worked. But when it came time for a break or lunch, he was gone. Many, but not all others exhibited the same behavior. Only two operators asked me when the time came for a break to come with them and continue the conversations, although many of those who did not would catch up with me later and pick up where we had left off. Thus talking was not viewed by operators as a burden or extra work. Taking notes was.

In the lab, of course, participants were presented their task as well as the treatment at once. From their standpoint, the two were indistinguishable. Those in the treatment group played a simulation game and maintained diaries of their ideas; it was all one task. The operators in the field study clearly distinguished between their work and what I was asking them to do. The fact that it was separate from their work meant that it was not necessarily required and, in fact, competed with their responsibilities for attention. The fact that it was extra raised the question in many operators' minds, "What do I get in return?" As I have noted, the returns from cooperating with me were uncertain.

*Explication required writing.*

The fact that the task involved writing raised a number of issues. Among the operators were several, perhaps nine, for whom English was a second language. Writing posed a challenge for them. One, on hearing my request and that he could take notes in his own language, said, "No. This will give me a chance to practice my English if you'll help me." However, he took no notes. Two expressed concern that their written English was not good. I tried to reassure both by saying that they could take the notes in whatever language they wanted and that if they did choose English it did not have to be good since we would review the notes together. Of these two, only one took notes (in English).

Although I never heard this from any operator, their manager told me of two complications created by asking them to write. The first was that the
native language of some was not really a written language\textsuperscript{2}. Thus, while they spoke English fairly well, they had very poor writing skills. There was another issue, one that he identified only very late in the project. Speaking, reading, and writing English were requirements for the operator job. Explication could reveal weaknesses in their writing that conceivably might threaten their position. They certainly would not feel comfortable writing in their own language under these circumstances. Thus an offer intended to take pressure off them actually may have raised some very difficult issues for a few operators and led to the avoidance that I observed.

*Explication is radically different from other things that operators do.*

I have already mentioned that operators were used to reporting only what they knew to be correct. Typically they did not diagnose problems. Explication called on them to offer hypothesis that would likely be wrong. It required diagnosis. Operators were used to directing people to engineers when they did not know something or did not want to address it. My field notes are littered with examples of operators telling me that I really should have been talking with the engineers instead of them. Explication demanded that they do the thinking. Operators had made a great deal of effort to structure their work. Explication was far less structured than the work they were used to. Potentially, it required skills that they had not developed. In short, this task put operators in very unfamiliar territory and represented a substantial change for them.

**Summary**

Most operators appeared comfortable talking with me. They seemed quite spontaneous, wanted to show me things, and told me about what was happening in the room. Talking to me seems to have become routine and uncontroversial. In the first part of the study, I never asked people to *do* anything or to make any decisions. I just talked with them.

\textsuperscript{2} The production manager is not a linguist. Nonetheless, he assured me that this statement is correct. Some operators' difficulty with writing was a property of growing up with their first language, rather than simply a matter of illiteracy as it is usually defined.
But my request to do more than just talk was a different matter. It was a change for them and one with implications on the many dimensions described above. It called for operators to take initiative in ways with which they were unfamiliar and uncomfortable. It challenged role boundaries in an organization where those boundaries were watched carefully. It singled people out and created enormous uncertainty for them. The task represented extra work and, because it required writing, was difficult or threatening for some people. Perhaps most importantly, it introduced change and, in the words of one operator, "[The] truth is, we're all in a rut. We just don't like change." In light of the interpretations of the request and task discussed here, the operators' responses to it become understandable, if not quite sensible. These interpretations of the request and of the explication task itself help explain why many operators' motivation to explicate was at least limited, often non-existent, and in cases even negative.

The third refinement, but one that must be kept separate for reasons I will discuss shortly, is the identification of issues surrounding explication. These issues were raised by the particular meanings of the request and of the explication task itself that emerged in the setting.

Although I could treat the issues discussed in this section as motivational barriers and integrate them into the framework presented in Chapter 4, I choose not to. The reason is straightforward; the two levels of meaning are implementation-specific and affect noticing only indirectly via explication. I do not propose that explication is the only way to affect these barriers and enablers and therefore do not wish to argue that the issues discussed in the meaning section are necessarily elements of any theory of noticing. Instead, they are parts of a theory of explication. Given that explication is the proposed route to noticing, they play an important practical role but their effect on noticing is necessarily indirect. In contrast, the barriers in the framework represent fundamental mechanisms and they have a direct effect on noticing.

5.4.2. The Intervention and Barriers to Noticing

The prior subsection focused on the intervention and its meaning to the study participants. In this section, I call on the framework developed in
Chapter 4 to analyze the operators' search behavior and results. In doing this, I identify several new barriers, not found in the analysis of the experiment, and offer several refinements to the framework that were not indicated by the experimental results.

The Motivation to Search

Task Motivation:

It is important to recognize that operators conceived of their primary task (that is, their jobs) quite differently. While a few felt that identifying and diagnosing problems, learning more about the production process, and therefore processes of search were within the scope of their job, many were so procedurally oriented that they did not see search-related tasks as included. These rule-focused operators saw any non-procedural activities, and thus search, as the engineers' responsibility. Thus, in this setting, unlike the laboratory, one must be careful to distinguish between task motivation with regard to their job and motivation to search. Operators who were motivated to perform routine tasks may have done so reliably but had no interest, either intrinsic or induced by external rewards, in search. One of these operators told me, "I do my best. But when it comes to solving problems, that's the engineer's job." Another said, "Look, I basically want to be left alone, do what I'm told, take home my paycheck."

Moreover, there are really two levels of search to consider. The first is the search for "problems" or, as I described it in the instructions for the intervention, "problems that affect the performance of the line or the quality of the product." The second is the search for their causes. Both types of search can involve noticing new variables, since problems' earliest symptoms are often different and distant from observable data on performance or quality. Not only did operators differ on whether they were responsible for non-procedural, search related activities, but also those who did step beyond the routine disagreed on which of these two activities – problem recognition and diagnosis – fell on their shoulders.

Predictably, intrinsic interest in the equipment and process appeared decisive in how expansive were operators' views of their role. One who was particularly motivated and was involved in both problem identification and
diagnosis told me, "I just know how things work. And when I don't, I want to. I've been fixing cars since I was a kid. This is the same thing." Others just told me that they thought the robots were interesting and that they wished they knew more about them. In a few cases, these statements were backed up by efforts to take courses on computer programming or mathematics. Those who expressed curiosity in discussions with me were the same ones whom engineers identified as most likely to report problems and to provide useful information (i.e., information beyond the problems' mere existence).

Another element in determining operators' interest in the search process was the rewards they perceived might stem from it. More specifically, only three operators' felt that they had been rewarded for good diagnosis work or for recognizing and reporting problems. When I asked what the rewards were, they reported that they were subtle: better relationships with the engineers, more respect from them. It is not clear that these rewards would be enough if they were not intrinsically interested in the first place. Several operators expressed that the only outcome was an increase in their workloads; they appeared not to value their relationships with the engineers:

"If I do help out. I never get the credit. No one ever says, Ellen helped me with this. Others don't help out. [The supervisor] can't count on them. So why should I help him out? If you care, they'll use you... Why does he [an engineer] ask me [to do things]? I asked him once. He said, 'Because I can depend on you.' I'm tired of being depended on."

These remarks came from the same person who, after explaining how she spotted a problem that others had missed, told me with some degree of pride, "That's why they like me on the line; I'm nosy." She was one of only two operators that I observed recognizing problems and addressing them before the robots began stopping. Thus she was a very capable, observant, and motivated operator. Nonetheless, her feelings about rewards, or lack thereof, tempered her interest in search. These feelings sometimes kept her, and many other operators from engaging in what one considered "extracurricular activities like helping the engineers."

Giving up

Based on my observations from the experiment, I defined giving up as an inference on the searcher's part that search will prove fruitless. I noted that
giving up was not necessarily the product of frustration but that successful search and problem framing could conspire to terminate search prematurely. I also noted that giving up could be a rational response to a situation, as well. These same themes emerge in the field.

Operators only rarely indicated that there was nothing to be noticed. They freely admitted that there were many variables associated with the equipment and the process that only the engineers knew: "You should talk with the engineers. They know everything. Much more than we do." Typically, however, they were referring to the variables that one might find in diagnosing a problem. For many operators there were only two relevant variables for spotting problems: robot stops and "falloff", subassemblies that failed the quality tests when inspected.

Indeed, most operators waited for the robots to stop before they would investigate. "Now, they're supposed to be scanning the line all the time," said one of the engineers (who had moved up from his original role as an operator). "You see how Lou is always scanning the line. Sometimes he picks up on things before they start seeing red lights [which indicate that a robot has stopped]. Waiting for red lights is a problem... There's not enough proaction on the operators' part." My field notes confirm this engineer's thoughts by revealing my surprise the first time I observed an operator scanning the line effectively:

He was working on the back side of the line [a position that involves a lot of loading and unloading of parts from the system], shuttling around observing, moving a stuck pallet here, watching what is going on at some station here, clearing springs out of a plastic parts bin... Surprisingly, much of what he did was not a response to a red light, unlike what most operators do, nor was it routine material handling.

Although this operator was clearly searching for problems, most others failed to scan the line in this way; instead, they waited for buzzers and red light to draw their attention. At that point, they would investigate, perhaps even do some diagnosis. From these operators' point of view, the absence of a red light meant more than just "no need to investigate"; it indicated that search was useless for the time being because there was nothing to find.
Chapter 5

The notion that search would produce results only at certain times arose in another way, as well. There were certain tasks during which operators felt that they were unable to notice problems or their causes. Some of these conclusions seemed quite surprising to me. For example, one operator repairing failed subassemblies told me that she "didn't see much when she was doing repairs." Another told me that "I'm not going to spot anything while I'm facing the wall," meaning that he would not find any problems to write about while doing repairs. I asked if the failed subassemblies wouldn't help identify some new problems or provide clues, he replied, "Like what?" These were not uncooperative subjects; I had had enjoyable, free flowing, and productive conversations with both. Furthermore, they were not asserting that search was somehow outside the scope of their task; they believed they could not identify problems while performing this task. This response seems particularly bizarre given that their task was to diagnose and repair failed product. They had the opportunity to see virtually every quality-related failure that occurred and thus were in an ideal position to identify problems and begin speculating on their causes. I observed similar responses from operators who were moving subassemblies and parts around the room. One asked, "what am I going to notice while I'm doing this? I'll be back on the line the day after tomorrow." In my time in the factory, two problems had been traced to material handling, one of which had led them to reconfigure every cart they used for moving subassemblies between lines.

For some operators, giving up had elements of an economically rational decision calculus. In discussing the intervention, one operator wanted to know what good identifying problems and their causes would do. She raised one problem as an example: "They'll never do anything about it... It's a complete waste. Nothing's happened in a year and nothing ever will." Other operators echoed this emotion in general rather than in relation to any specific problem. In short they were asking, "Why should we diagnose problems when experience shows that no one is going to do anything about them?" This should not be confused with motivational concerns related to explication or even with participating in the study. Here, the operators' issues were with search and the lack of meaningful results that could be expected from it.
Perceived need for improvement

In the experiment, I noted that participants paid attention to different variables for the purpose of interpreting feedback. The particular variable and, most importantly, the participants' expectations for the values it would take on proved critical in determining the degree to which they felt a need to do better and were thus motivated to search. Once again, the field proves more complex.

As I have already mentioned, operators differed quite fundamentally in their interpretations of their own roles. For those who saw their role as executing procedures, the lines' performance was the engineers' problem. Poor performance was no cause for search because of their particular conception of their jobs.

Operators could use several indicators of performance. There were daily production figures and reports specifying defect rates. They witnessed robot stops and thus sensed automatically their frequency. As in the experiment, one needs to examine the expectations that operators formed around each of these performance indicators. Quality had been steadily improving over time and defects were at a very low rate. If anything happened to product quality, operators generally sensed the problem immediately; the physical presence of rejected subassemblies was a highly salient cue that prompted action. More subtle changes in quality were likely to be picked up at the weekly engineering meetings to discuss each line's operation.

The operators were aware of the daily production figures (each robot had a panel display that showed how many units had been produced during the shift) and they could observe day-to-day fluctuation, which was high. As long as production did not drop below schedule, however, the fluctuations were not salient. In fact, two of the lines' performance was well below their peak levels. Only two operators mentioned this fact in talking with me. Schedule rather than past history was the salient expectation for operators; only two stopped to wonder why they could not routinely (or ever) meet schedule within the first seven hours of the shift when that should have been possible. It is important to recognize that the production schedule had been reduced, as well, making it an easy target to meet in an eight-hour shift. It was only the
rare and dramatic problem that would lead to a failure to meet schedule. Schedule rather than past performance was the more salient cue, however, and expectations for performance slipped as the schedule was reduced.

Negative consequences to search

I have discussed in the section on the meaning of the intervention the climate in which the operators worked: the degree to which their roles were circumscribed and the negative social consequences to stepping out of role. Search activity, and particularly extensive interactions with engineers that would likely result from them, might be conspicuous to others and might lead to conflict. In the face of these disincentives, operators who were not intrinsically interested in search might find that it offered a net negative return.

Summary

Participants' levels of interest in search varied considerably. In part this was due to operators' different conceptions of their job. In part it was due to their level of intrinsic interest in the equipment and process. It also appeared that any extrinsic rewards were subtle and the risks salient. The intervention did not alter this balance. It appealed to those who exhibited great intrinsic interest and perhaps offered small extrinsic rewards associated with diversion or the opportunity to speak with someone new. But it did not alter the negative returns to search.

Many operators exhibited symptoms of giving up. Sometimes, it was a conditional kind of giving up, based on the perceived impossibility of noticing anything in certain roles. However, operators also gave up when they felt that the results from search would not be used and thus would ultimately not yield results. Finally, the poor performance of the lines was salient to very few operators; what caught their attention were catastrophic failures (like unplanned shut-downs) and failure to meet schedule. Since the schedule was not demanding, it was rare that operators received salient feedback indicating a need to search. It seems possible that had operators explicated, they might have challenged some of the assumptions underlying giving up. Given the focus of the intervention, first on problem identification and then on diagnosis, however, this result seems unlikely. I
propose that the intervention could not have produced search without first
directing operators to explore the assumptions driving motivation.
Furthermore, the one-on-one approach and volunteer nature introduced
motivational issues that even if made explicit could not be overcome without
establishing different extrinsic rewards for the activity.

**Scripts that Affect the Search Process**

In the field study, several script-based barriers emerged that were not
observed in the lab. I have alluded to several already, but will make them
explicit here. In addition, I describe how several barriers, familiar from the
laboratory, manifested themselves in the field.

The new barriers are scripts offering alternatives to search, specifying
when to search, and specifying when to act on the basis of tentative
inferences.

*Scripts offering alternatives to search*

When problems were identified or when potentially relevant symptoms
such as robot stops occurred, many operators used scripts that were
alternatives to diagnostic search. These were: ignore and report.

When problems were severe, that is, they shut down the line, operators
would call engineers to come and help. When the problems were less severe,
they rarely did this. Instead they would ignore the problem until either it got
worse or until an engineer happened to walk by. At that point, the operator
might get the engineer's attention and let him or her know that, for instance,
a given robot was stopping a lot. As described above, operators rarely
provided diagnostic information, largely – but not entirely, as we shall see
below in "Scripts specifying when to act on the basis of search" – because they
had not attempted any diagnosis earlier.

In short, their scripted response to problems was: If it's bad report it
immediately – if necessary, by telephone. If it's not bad, wait until it gets bad
or an engineer arrives. Do not make any special effort to diagnose it, tell
other operators, or maintain a list of potential problems to share with
engineers.
Of course, such a script represents an extreme and there were clear exceptions. However, their primary effect was to highlight in my mind what operators were capable of and how rarely it happened. The situation was not just a function of operators' abilities. Their behavior was guided by scripts that had become the norm.

Scripts specifying when to search for problems

In discussing motivation, I mentioned how operators essentially gave up when they rotated through different roles. It is important to recognize that one operator's giving up, when communicated, might become a less experienced operator's script.

Perhaps the most damaging script of this type was one that dictated search for problems only when there was a robot stop. The most experienced and motivated operators did not use this script; even when they did not walk up and down the line, and thus appeared to be using it, in fact they were not. I observed one who would remain rooted to one spot, moving away from it only occasionally except to address stops. He told me, "I like to stand here because it's the only place where I can see everything that happens on the line." But other operators clearly did not watch in this fashion. Rather than scan the line, they waited for buzzers and lights.

By the time a buzzer sounded and the lights flashed, the robot would have stopped. Unless operators restarted the machine and watched it until it stopped again, they would be unlikely to identify any process-related causes or symptoms of the potential problem. With only three exceptions that I observed, operators did not exhibit this behavior unless asked to do so by an engineer. Instead, they would watch two or three machine cycles (insufficient given the relatively low probability of a stop occurring in any given cycle) and then move on.

Scripts specifying when to act on the basis of tentative inference

An important script that guided operators' behavior specifies inaction until confident that one is right. This script applied to reporting problems. One operator described a problem for which she had summoned an engineer. "Then, when he shows up it doesn't happen again... So the machine just makes me look stupid." The end result was her reluctance to report problems.
until they occurred consistently and frequently; much later than makes sense. Another operator put it simply, "I don't open my mouth unless I'm sure I'm right."

They interpreted engineers' lack of response as feedback indicating either that they were wrong about something being a problem or a lack of caring on the part of the engineers. Both interpretations made operators reluctant to respond until problems became really serious. Unfortunately, the engineers' slow response was in part the result of a heavy workload and factors such as waiting for special parts. In one case, for example, the fabrication and testing of a component for a robot took more than three months; the delay was cited by numerous operators as an example of the engineers' lack of caring.

The operators who explicated reported in their notebooks only problems with grave consequences, despite the fact that there were many others that occurred. When I asked why they reported only these problems one said that he "wanted to make sure that [he] had found a real problem to talk about." Another wanted to make sure that they "were important enough for [me]." Furthermore, they did not speculate on causes. They never reported multiple causes or alternatives to their ideas. They wrote down and told me only those things of which they were sure. The scripts that dictated action in their work were carried over to the explication task.

Learning based on small numbers of observations

An important series of scripts in the experiment were those related to processing based on small numbers of observations. In the experiment, I observed a willingness to use very small numbers of observations both in generating hypotheses and in testing them. In hypothesis testing, such reasoning was constructive, as is made clear by the case of the single participant who did not do it. In hypothesis testing, however, it led to rigid use of incorrect inferences and premature rejections of correct ones.

The operators' reluctance to act on the basis of anything less than near-complete confidence meant that they could not respond to small numbers of observations. Since they represented the organization's eyes and ears, this meant that hypothesis generation in any form leading to organizational action could not take place on the basis of small numbers of observations.
Problem identification remained stalled at hypothesis generation in much the same way that it was for the single participant in the experiment who would not experiment until he observed conclusive evidence disconfirming the linear price-quantity model.

There was very little opportunity to observe hypothesis testing among the operators, in part because of the role constraints discussed above. The constraints within which operators worked and thought appear to be more the product of scripts and frames than the result of conclusions made on the basis of hypothesis tests of too few observations.

Scripts for responding to disconfirmed hypotheses

In the lab, we observed operators who dropped variables because hypotheses containing them were disconfirmed. On the shop floor, operator's roles generally prevented them from taking action to solve problems. Thus, what information they did offer (i.e., their reports of problems) could only be disconfirmed by others, often engineers. Thus, disconfirmation was a public event. In a climate where being right was important, disconfirmation often seemed to lead to withdrawal from search altogether. Many of the operators reported that they felt the engineers did not respect them and made them feel "stupid", to mention a word used by several. The message that many operators took away from any instance in which they reported problems and were in some way wrong (i.e., disconfirmation) was that they should not participate in any aspect of the problem solving process.

Scripts restricting the range of experimentation

The range of actions that operators could take to correct problems or even to diagnose them was restricted severely. They could not, for example, adjust the equipment in any way or switch from one lot of parts to another on an experimental basis. Although there was a certain amount of fiddling that operators would do to the equipment, typically it only took the form of prodding devices for supplying parts to the robots that occasionally jammed or nudging trays on conveyors that got snagged as they moved along. For the most part operators were observers. Experimentation analogous to that of the participants in the experiment was impossible.
Chapter 5

Thus operators' only means of "experimenting" was to report their ideas to engineers; their primary feedback came in the form of other people's responses. Given the public nature of these exchanges, operators restricted their "experimentation"—that is their offering of ideas—to those in which they were highly confident. This restriction on experimentation more closely analogous to requiring large numbers of observations for hypothesis generation than it is to the restricted experimentation by subjects in the prior study.

In sum, the operators' experimentation was limited by the boundaries of their roles and by their unwillingness to offer hypotheses. Thus the issue of a limited range of experimentation analogous to those observed in the laboratory did not arise. One can imagine that if operators were given some room to experiment, they might exhibit similar conservatism. However, this remains a matter of pure speculation.

Scripts for examining and responding to one's own frames and scripts

In Chapter 4, I noted that the ability to draw conclusions and yet remain skeptical seemed important for the noticing process. I observed that successful participants in the experiment seemed to have been able to question their assumptions and actually did so quite consciously.

I observed nothing even remotely resembling this behavior among operators. Even when I questioned assumptions that emerged in conversations—for example about engineers' lack of caring or their lack of appreciation of the operators, the permanence of an accounting system with disastrous side-effects for interdepartmental cooperation, or the restrictions that they perceived in their roles—they exhibited a near-universal unwillingness to reconsider their viewpoints.

There are at least two reasons that might explain this, including a lack of openness with me on the part of the operators and the absence of circumstances that would trigger use of such scripts. Given the operators' general willingness to talk with me, even about controversial issues, I doubt the first. The second is a more likely candidate. Given that operators did not sense a crisis (i.e., the equivalent of the frustration that some participants in the experiment experienced), it makes sense that they would not exhibit these
scripts. However, it is interesting that they were unable to respond to the inquiries mentioned above with anything other than stubbornness.

*Role related scripts*

In the experiment, I discussed (and interpreted) scripts as though they were an inherent component of human information processing. However, the scripts discussed here should make clear how dependent are people's judgment heuristics on scripts identified with their roles and supplied to them by the organization around them.

I noted that participants in the experiment sometimes framed the problem differently and thus had somewhat different goals (e.g., some viewed the task as beating a program that I had created while others did not try to "psych out" the game). I noted how these different framing invoked different modes of evaluating the feedback they received. One can think of these different frames as the participants' views of their role and the accompanying scripts as role-related scripts. In general, however, the participants in the experiment agreed on the basics of their task: they were all trying to discover how to make decisions better in order to get good scores.

By contrast, the field provided participants with fundamentally different conceptions of their role, particularly in relation to search— in the forms of both problem identification and diagnosis. For some operators, search was central to their conception of their role. But to others it was secondary and to a third group it fell outside of their role and was clearly somebody else's job.

The scripts discussed in this section reflect those different conceptions of role, are an essential part of the construct of role, and are maintained institutionally (as we have seen) through interaction with others.

Some operators, then, had scripts that essentially prohibited the very activity that I was trying to elicit through the intervention. The intervention did not draw attention to these role-specific scripts and provided no incentive or sanction to reconsider them. Given this, it comes as no surprise that operators did not challenge them as would have been required for the full participation of many.
Chapter 5

Frames that Affect the Direction of Search

As with participants in the experiment, operators exhibited frames that directed search. Generally speaking, the frames that emerged in the field direct attention to variables, rather than away from them.

In discussing frames, I need to distinguish once again between the search for problems and the search for causes. Frames determined what was considered a problem and what kinds of variables might be symptoms of problems. They also affected the search for causes. I discuss both in the subsections below.

Frames supplied by the context

In the laboratory context, the case and instructions provided a context that affected some participants’ framing. The frame that appeared most commonly was one that excluded specific variables (i.e., the advertising). In the field, context appears to have played a greater role in directing attention to specific variables than excluding them from consideration.

It is quite clear that the vast majority of operators used two cues to identify problems: stops and the presence of subassemblies rejected during inspection. Both these measures were used in printed reports. They were discussed in meetings. They were the first two measures to which I was introduced as I was taught about the process. In addition, stops were highly salient because of the red lights and buzzers, and also because they demanded action from operators. If a given stop occurred frequently enough to "give me a headache", in one operator's words, then it was a problem. Similarly, rejected subassemblies would literally pile up next to the line, also a highly salient cue. Typically, whatever defect rose to the top of the list or resulted in a large pile of rejected subassemblies was addressed as a problem. The point at which a stop became "a headache" however was quite subjective. The point, however, is not the subjectivity of assessment but the presence and continual reinforcement of the frame specifying these two variables.

I observed another common frame used not to identify problems but to describe them; although rarely employed spontaneously, the operators universally were aware of it. When a stop became frequent, engineers would sometimes ask operators to record certain information each time the stop
occurred (I observed this four times). It happened frequently enough that operators occasionally would do it spontaneously (I observed this once). The two variables were the number of the platen (the identification numbers of the platforms on which subassemblies moved through the system) and the position on the platen (since each platen held several subassemblies). This descriptive information was helpful for identifying causes such as defective platens (if the stop occurred only on one) or ruling out others such as bad parts (if the stop occurred only in one position). If operators took the initiative to diagnose a problem, this information would likely be some of the first they would collect.

There were generally loose parts lying around in odd places. When operators cleared stops, they would almost always take a look for loose parts that were in the way. It was perhaps the one cause that operators universally checked for. However, their diagnosis did not go very far because I never heard them mention loose parts as a problem. They expected to clean them up at the end of the week or when necessary because they obviously caused some stop; with perhaps two exceptions, they appear also to have failed to recognize them as useful symptoms. In short, the variable had been distinguished, was linked into a model of stops as a cause, but did not function as a symptom requiring diagnosis.

The field work confirms that context can supply people with frames. In the experiment, I did not even mention the implicit model linking price and quantity. Here, the base model was more obviously supplied, hence the salience of frames suggesting rather than excluding variables. Perhaps more interesting is that observations in the field suggest that frames not only specify variables; they also specify role: either symptom or cause. A variable framed as a cause might not be recognized as a symptom to investigate or report on its own.

Frames that fix the set of candidate variables

In the experiment, participants appeared at times overwhelmed with complexity and responded by consciously deciding to consider only the variables that they already knew about, at least temporarily. In this context, I would argue that role-related scripts limiting search were dominant; I observed no clear frames fixing the set of candidate variables. Operators did
not feel that they could not cope with the variables they knew; they felt it was someone else’s responsibility to do so. Those who viewed their role differently exhibited no frames limiting them in this way.

Frames excluding what you cannot control

Lack of control was intimately related to giving up. As I have mentioned, there were problems whose causes appeared to lie in the behavior of other departments in the organization. One operator doing repairs told me that he knew the subassemblies were fine when they went out so the problems with the returned ones were caused by something happening in another department. However, he also told me that it was not worth thinking about because, "...you don't understand. They aren't going to change." This example might represent a script: ignore problems that appear to be caused by things you (or we in this department) cannot control. On the other hand, such thinking might translate into a frame directing attention away from anything relating to the behavior of other departments or factors not directly under the operators' control. It may be significant that this example represents the only instance when operators discussed anything that happened outside their department (including the parts vendors, whom the engineers mentioned on many occasions).

Frames constraining the form of relevant data

In the experiment, one of the difficulties participants had in noticing the advertising variables was distinguishing textual variables in a problem that was otherwise dominated by numeric information. I suggested that the participants had frames that directed their attention toward numeric information. In the field, neat categories relating to the presentation of data do not emerge. Variables did not appear on the printed page or the computer screen but rather as a diverse set of behaviors and phenomena, not so easily categorized.

In the field, the only types I could distinguish related to topic areas. In describing problems from their notebooks, operators generally discussed two types of causes: variables relating to parts (e.g., the presence of flashing at the mold seams of plastic parts or bent springs) and variables relating to the equipment (e.g., sensors failures, calibration problems, a loose part that had
fallen into the wrong place). Often these "causes" had deeper root-causes that were related to people's behavior. Although engineers mentioned them, neither in their notebooks nor verbally did operators tell me about such causes – the quite real contributions that people made to problems, for example, through poor maintenance, improper handling of parts or subassemblies, or failure to keep the equipment free of loose parts. One might attribute this to an unwillingness to talk about other people; however, this is inconsistent with operators' willingness to criticize others when not discussing the causes of problems (e.g., this person gets special attention, this person does not deserve the rank she has, this person always gets to take his lunch when he wants, etc.).

It is possible that operators interpreted my interest in problems relating to output and quality as an interest in the domain of "technical problems." Such an interpretation is consistent with several operators' statements that I really should be talking with the engineers. The types of symptoms and causes reported are associated with that domain whereas some of the ones mentioned above perhaps are not (although engineers did identify them). Given the absence of discussion of this type of variable, it seems possible that the operators have frames that direct their attention to other matters, either as problem or as causes.

*Frames based on specific interpretations*

In the experiment, participants sometimes interpreted the advertising variables without even looking at them. These conclusions often went unchallenged. Certainly, operators did not challenge their assumptions that any variable outside their control could not be changed. Otherwise, I observed only one such filter.

There was a class of problems that operators did not report to me that, in their opinion, could not be solved. These emerged only in the last week of the study because I began asking about symptoms that I observed. In several cases, the operators had given up on these problems; they had become the status quo and were no longer reported. I observed that in some cases the assumption that operators made was not necessarily valid and that their failure to report a problem actually prevented engineers from getting the
feedback that they needed to address problems that in fact could be solved and justified attention.

Summary

The data and interpretations presented above are useful both in extending and refining the framework presented in Chapter 4. The extensions are straightforward, new barriers or enablers identified in the field work. Four new types of barriers emerge from the field study: disincentives to search that can dominate intrinsic task motivation, giving up based on the fact that the results of search will not be used (rather than an analysis only of knowledge gained versus cost to get it), scripts that provide alternatives to search, and scripts specifying when to search.

In refining the framework, three points come to mind. The first is that we need to distinguish between types of barriers and the sources that produce them. It is useful, for example, to distinguish between scripts that are role-related and other types (e.g., those that appear to be provided as an innate part of our human information processing apparatus); however, it is important to recognize that scripts can also be classified on the basis of their behavioral consequences, independent of their source. The classification scheme in Chapter 4 confounded the two categorization schemes (e.g., frames fixing the candidate set might have been supplied by the context and yet these are listed as two types of frames). An outcome of the field study is recognition of two distinct dimensions for classifying frames and scripts: source and behavioral consequence.

A second refinement is the recognition that the world (and our models) consist of both causes and effects, and that frames may apply not only to variables but to variables in specific roles (e.g., as symptoms or as causes). The laboratory study did not reveal variables framed both as cause in one model and effect in another. Examples from the field, however, suggest that the frames applying to a variable as cause, might be very different from those used when it is treated as a symptom instead, and might lead to different search behavior.

Finally, many of the frames observed in the field directed attention to, rather than away from, specific types of variables; this fact should remind us
that, although frames do act as barriers, they also can play an important enabling role under certain conditions. Those circumstances include: complexity, which frames can help reduce thereby making problems tractable, and stability, which makes frames' effects as barriers unimportant. Slow change, however, may prove to be the most challenging situation. While slow change does create the need to notice, because it provides a stream of disconfirming feedback that grows only slowly over time, it is also likely to induce motivation-based barriers to noticing that halt search.

5.5. Summary and Conclusions

This chapter has used the field study as an opportunity to develop a set of issues relevant to implementing successful explication-inducing practices, essentially a set of barriers to explication. Because these barriers act on noticing only indirectly, via explication, and because explication is not necessarily the only intervention that might affect the mechanisms that enable or block noticing, they constitute a related but separate theory, a partial and tentative theory of explication.

The field study has also helped identify several new barriers that I propose have a direct affect on noticing and thus should be included in the original barriers framework developed in Chapter 4. The field work also led to observations useful in refining the framework.

With these refinements in mind, I will review the framework briefly and include as a coda the explication-specific issues identified at the start of this chapter. I conclude the chapter by offering an alternative intervention that addresses some of the issues found in the field setting.

5.5.1. The Two Frameworks

Barriers to and Enablers of Noticing

The framework consists of a series of conditions that enable or disable the noticing process. They fall into three broad groups: those relating to the motivation to search, those that are a consequence of scripts, and those that are a consequence of frames. For noticing to occur, people must be motivated to search, must use scripts that enable or at least do not block search, and must
not have their view of the potentially relevant variables obscured by their
framing of the problem.

Motivation to search

Perhaps the most fundamental motivational enabler is inherent interest
in the phenomenon in question. Without it search is unlikely at best. As
Amabile (1988) notes, inherent interest is perhaps the most fundamental
requirement for creativity, as well. Inherent interest, however, can be
overcome by people's conceptions of their role, if they happen to exclude
search, or the presence of negative incentives to search. Thus, there are a
range of factors relating to individual, task, and context that determine if that
person will be motivated at all with regard to search. Beyond this base level,
there are the twin barriers of giving up and low perceived need to search.
Either can halt the search process and both are critically related to frames
which are often untested and untestable, and therefore little more than a
matter of faith.

Scripts specifying search activities

Script related barriers fall into the following categories. There are those
that enable and disable the search for variables. These specify when to search
and alternative courses of action, such as reporting or ignoring. In the
experiment, scripts directing to examine and test their assumptions appeared
to enable noticing and although not observed in the field (which is consistent
with the low apparent rate of noticing among operators), I think it likely that
they do play a role. Three more scripts were identified, those related to the
processing of small numbers of observations, those related to the treatment of
variables that were part of disconfirmed hypotheses, and those that limited
the range within which people experiment with known variables.

It appears that some scripts are supplied through people's roles,
communicated to them and reinforced by their interactions with others.
Another set of scripts, those related to hypothesis generation and testing,
appear to be either the product of instinct (i.e., some basic nature of the
human brain) or culturally-learned.
Frames that direct attention during search

The final type of barrier takes its form as frames that relate more specifically to variables than to the activities of search. These frames can exclude variables or include them; often, however, by including some they act to exclude others. They can be used to constrain a problem by fixing the candidate set. They can drive exclusion of variables, for example those that we do not control, or restrict search to certain types of variables. The specific interpretations identified as frames in the experiment are significant in that they demonstrate the ease with which people form frames and the effects they can have on search even when they are essentially unsupported by evidence. Low confidence in frames appears to have little effect on their power to direct search.

Like scripts, frames can be supplied by the context (e.g., statements by other people, rules) and reinforced (via reporting protocols, for example) or they may be constructed on an ad hoc basis by the individual.

Barriers to and Enablers of Explication

The field work has revealed a series of barriers that affect noticing only indirectly, via their restrictive effect on explication. They fall into two categories, those that were the product of the intervention request, but not of the task participants were asked to do, and those that were a response to the task of explication itself.

Any intervention introduces change. In this case, just requesting operators to participate presented issues for them. It called for initiative on their part and challenged existing role boundaries. It asked them to cope with uncertainty which, for the most part, was seen as offering risk rather than opportunity. Finally, it singled people out. Thus, even before operators knew what I wanted them to do, there were barriers to their participation.

The specific task also caused problems. First, it constituted extra work. The word extra is important; the task was seen as separate from their regular tasks and completely outside their role. Therefore it competed for attention and was at best a second priority despite the limited time involved. It called for writing and thus was challenging and even threatening for some because it might point out skills they were supposed to have. Finally, explication was
a task quite different from what operators were used to. At one point, my field notes read, "It's going to take a lot of coaching to get them to do more than just document what they think they know." I think that insight was correct. Explication called for skills that perhaps operators did not have or, more likely, that they were not used to using. Furthermore, in comparison to their usual tasks, explication was quite unstructured. Given the operators' past resistance to poorly structured work, their avoidance of explication even when they volunteered to participate is understandable.

These barriers tell a familiar story about introducing change into organizations and reinforces points made by others in the past (e.g., Beckhard & Harris, 1987; Beer, 1980; Schein, 1969, 1987). In identifying them, however, this work represents an initial step on the road toward designing more effective interventions in the future. Perhaps more importantly, however, this work begins the task of building a theory of explication and describing more fully the relatively sterile construct presented in the earlier chapters of this document.

5.5.2. An Alternative Intervention

In closing, I would like to offer an alternative intervention that addresses some of the barriers identified in the chapter. There are additional steps beyond the intervention that I would recommend to develop the operators' basic skills. Here, however, I focus just on the intervention itself. The alternative intervention would:

- Be integrated with the work
- Be supported by a philosophy and clear goals
- Include coaching so operators would have the opportunity to develop the skills they need
- Be intense enough to alter scripts and frames that are not otherwise reinforced
- Provide social rewards, particularly in terms of respect and acknowledgment of expertise or contribution
- Reward effort on an ongoing basis
- Allow for personal growth and advancement
Chapter 5

In this setting, I think the leverage point is in the relationship between engineers and operators. Explication would have to be implemented as part of that relationship. Although demanding of the engineers' time, such a program would offer benefits for the engineers. If constructed properly, it could offload work from them.

First, the operators' role as diagnostician would have to be clearly established as part of their task. Given this setting and the operators' efforts to limit their work to the well structured, diagnostic work would have to be rewarded through some kind of incentive system. For operators who were intrinsically interested in the operation, increased autonomy, involvement in problem solving, training, and promotion would be appropriate rewards. For those who were not intrinsically interested in the task, no reward would result in true explication or good diagnostic work.

The relationships between engineers and operators would be critical to the success of the intervention. Both would need coaching to facilitate the process. First, they would need to overcome the barriers to communication already present in the environment and hinted at earlier in this chapter (e.g., the engineers don't care, they don't listen, etc.). Although it introduces the problem of singling operators out, this work might be done best on a one-to-one basis. Perhaps one approach would be to plan to put all operators through it and select the sequence on a lottery basis.

The intervention would call for operators to identify problems and then to explicate with regard to their causes. Perhaps they might work in pairs, developing their ideas before bringing them to an engineer. The engineer's first responsibility would not be to start shooting down the operator's models but to create more alternatives. At that point, the operator might work through the possibilities, narrowing down the diagnosis until he or she could go no further. Additional diagnosis and problem-solving would be done jointly; the operators would remain involved, at least aware of the process and the decisions being made.

The engineers play a very important role in the process. First, they must counteract giving up and any perceived satisfaction with performance by helping operators see the arbitrariness inherent in any such evaluation.
Second, they must watch for scripts that interfere with search, point them out to operators, and provide alternatives. Finally, they must try to identify frames that restrict the operators' view and encourage operators to uncover them for themselves and to question them.

For the engineers, it is a difficult and time consuming task. The use of an outside facilitator might offload work from engineers and provide the support they would need in their new role. The supervisor or production manager might act in this capacity and thus be able to evaluate the efforts first-hand.

This sketch of an intervention begins to tackle the most difficult barriers to noticing and could counter many of the barriers to explication that I observed. It is very demanding on someone in a facilitation role; but I do not see a less intense alternative being effective given the types of changes that are required. Furthermore, I see the demands of facilitation dropping dramatically over time as the practices become institutionalized and operators can take on the bulk of responsibility themselves. It should be clear by now that the intervention used in this study would have been altered considerably if the framework developed in this and the previous chapter had existed at the time of its design. Since the intervention used was necessarily exploratory, in a sense a pilot, the learning that it has produced is quite satisfying. The principles discovered here will lead to new, more refined interventions.

5.5.3. A Note on Organizational Culture

It is very important to note that the firm I studied is known in the business community for its forward-thinking practices, employee ownership programs, pay-for-learning systems, and use of TQM. The business press has praised the firms' practices and academics have written about its policies. Apparently, neither group has examined the long-term results or looked at this factory's operations in depth.

The factory's history is filled with efforts on management's part to raise worker skills, involvement, and autonomy. The firm's tuition reimbursement policy is as generous as any in industry, and teachers from local universities, technical schools, and high schools teach courses on site.
Courses that are work-related, specifically those relating to process control, are taught on company time. Every employee with whom I interacted had been through at least introductory training in TQM. They all, for example, had received instruction in the use of control charts. However, these programs and policies have failed to address the firm's underlying cultural problems, many of which should be apparent from reading this chapter. The end-result is widespread apathy and disillusionment with change efforts (for more, see Beckhard & Harris, 1987; Beer, 1980; Schein, 1969, 1987).

Thus, beyond its lessons about explication, this chapter carries another, implicit message. It offers a warning about inappropriate and incomplete solutions to organizational problems. Using training focused primarily on skills to address organizational problems that are cultural in nature will fail. Even when new skills are needed, training programs that do not simultaneously address cultural barriers – which generate resistance to change and ensure that newly acquired skills never get used – waste resources and compound problems rather than solving them.
6. Conclusion

6.1. Recapitulation

I began this dissertation with the observation that today's business climate requires people in organizations to notice new variables if their firms are to survive. Noticing is both the stimulus for innovation and a fundamental component of the process of adaptation. Throughout most organizations there is a general shortage of noticing relative to need. Yet noticing, as defined in this work, has received little attention in past research. Therefore, it is important to study the variables that affect noticing, in particular those that might imply techniques for managing it.

Chapter 2 started with a definition of noticing. It presented a model of the sequential states through which variables progress as people construct mental models. A variable is considered noticed if it is part of a specific model's "candidate set" (i.e., the set of variables actively considered potentially relevant to the model). This definition implies that people may be aware of variables and yet not have noticed them in the context of a given mental model. It also implies that variables may be "unnoticed" if they are forgotten, framed out of existence, or come to be viewed as irrelevant. Finally the definition makes clear that there are two steps on the path to noticing. The first is distinguishing the variable, that is creating it as a mental construct that represents some property or combination of properties of the world. Once a variable has been distinguished, the second step is to hypothesize its relevance to some particular model, thereby introducing it into the model's candidate set. There are additional steps in the process of model construction; however these go beyond noticing.

The vast bulk of work related to noticing actually excludes the distinguishing of new variables. The exceptions (e.g., Hull, 1920; Klayman, 1988; Klayman & Ha, 1989; Wason, 1960) have either focused on hypothesis testing or have used predictors that people cannot influence. Thus this work has not helped with the task of identifying variables that might be useful in managing the noticing process.
Based on pilot work, I proposed that explication – the process of writing down hypotheses and thereby making explicit one's mental models – can stimulate noticing. This hypothesis is consistent with literature related to explication; however such a link has not been proposed in the past. The hypothesis also introduces a variable, explication, that has potential to be managerially useful. This motivates the dissertation's first question:

Does the process of explication affect the thought process in such a way that noticing new variables becomes more likely?

The third chapter presented an experiment that addresses this question. Second-year Sloan Masters' students performed a repeated decision-making task under either of two conditions. The task called for them to make 40 sequential decisions, setting prices for a product and receiving feedback on how many orders they received. The treatment group was asked to maintain a written protocol of their hypotheses as they executed the task. The control group performed the task without maintaining the protocol; pens and paper were available to them. Noticing was measured by whether or not they looked at two variables that were available to them but not obvious, and either formed hypotheses about or tried using them.

The treatment increased the likelihood of noticing the non-obvious variables; 69% of the treatment group noticed compared to 24% of the control group. Further analysis yielded several additional results. First, the treatment was effective in altering participants' recording behavior. Specifically, although participants in both the treatment and control groups spontaneously recorded data, only treatment group members recorded their hypotheses about causality (i.e., explicated). Second, the results show that the treatment's effect on noticing is mediated by explication. Logistic regression analysis demonstrates that the treatment is no longer a useful predictor of noticing once explication is controlled. Fully 82% of the people who explicated noticed the non-obvious variables compared with only 23% of the non-explicators. Further analysis rules out the hypotheses that data-recording, time pressure, or effort acted as mediators – either on their own or together with explication – in the relationship between the treatment and noticing. Additional analyses suggest that explication's effect is not due merely to superficial effects such as slowing down people's natural thought
processes. Instead, explication seems to be effective in causing noticing because it changes people's thought processes in important and fundamental ways.

Having established a causal link between explication and noticing, I ask:

What are the mechanisms through which explication has its effect on noticing?

The fourth chapter analyzes in depth the behavior of the participants in the experiment. It proposes a model of the inference process, develops three sets of mechanisms affecting this process, proposes that these mechanisms are the means through which explication influences noticing, and demonstrates that these mechanisms, as a group, mediate the relationship between explication and noticing.

These mechanisms are termed barriers to and enablers of noticing, and fall into three groups. The first group influences the motivation to search. In the experiment, intrinsic interest emerged as a basic enabler of noticing. Two judgments, one about the need for improvement (and thus the need to search) and another about whether search will yield useful results (i.e., giving up), appeared as motivational barriers to noticing. Both these inferences were made using frames that were arbitrary and often went untested.

A second group manifested itself as scripts directing the search process. Several scripts governed the use of small numbers of observations in inference. One participant's insistence on large numbers of observations interfered with hypothesis generation. Other participants' use of small numbers of observations led to Type-I and Type-II errors in hypothesis testing; in the experiment, these errors often went uncorrected. A rare but particularly interesting script led participants to respond to disconfirmed hypotheses by dropping the associated predictor variable(s) from their candidate sets. The inference about the variable's lack of relevance can never be justified; the decision to drop it can be justified rationally, but only when the modeler cannot identify additional hypotheses that include it or when there are more promising variables available with which to work. Finally, participants exhibited scripts that supervised the model building process at a higher level; these scripts specified whether and when to search, and

- 219 -
managed the choice between expanding the candidate set (the search for new variables) and constructing better models using variables within the current candidate set (search within the candidate set).

The third group consists of frames that affect the types of variables people notice. In the experiment, frames generally constrained search. Participants exhibited frames specifying the types of variable to consider. For example, one frame directed participants' attention to numeric variables. This frame may be a specific instance of a more general one that specifies similarity between cause and effect, a cue to causality from Einhorn and Hogarth's (1982) framework. Frames directing attention to one type of variable also directed it away from others; thus participants expecting numeric variables failed to distinguish useful text-based variables. Frames also excluded variables. Participants sometimes framed any but the current candidate variables as necessarily irrelevant, thereby fixing the candidate set. Others excluded variables outside of their own control.

A hallmark of all the frame-based barriers (which include perceived need to search and giving up, motivational barriers based on problem frames) was the degree to which they were arbitrary, rigidly held once formed, and drove the search process even when unsupported by data. Given this, scripts that directed people to search for their assumptions, to see if they had misinterpreted something, and to reframe the problem appeared to be powerful enablers of noticing.

Although the fourth chapter relies primarily on qualitative analysis, it uses logistic regression to demonstrate that the barriers and enablers identified in the chapter do, in fact, mediate the relationship between explication and noticing. However, the analysis also demonstrates that they explain only a portion of explication's effect, suggesting that there are more barriers and enablers to be noticed and analyzed.

The chapter also outlines the role of data-recording in the inference process. It demonstrates that data-recording plays an important role in parameterizing models (i.e., search within the candidate set) and proposes that, because data-recording increases the accuracy of hypothesis testing and
reduces memory-related biases, it should influence some of the barriers and enablers, as well.

The fifth chapter opens with a discussion of the experiment's limits with respect to generalizability. Many arise from the laboratory's absence of realistic social context. Chapter 5 presents a field study whose goal was to identify additional variables that influence the explication-noticing relationship as it might occur in an organization. Specifically, it sought to answer the question:

What are the organizational characteristics likely to affect explication and moderate its relationship with noticing?

I studied a group of production operators in a manufacturing organization, first just observing, then introducing an intervention intended to produce explication. While explication was unproblematic in the laboratory, the field study revealed several barriers to explication – influences on explication and on whether the intervention was effective in inducing it – that arose out of the meanings that the request to participate and the explication task itself held for the operators. Both responding to the request and explicating challenged role boundaries that the operators had carefully constructed and maintained over time. Also, explication called for skills that were qualifications for the operators' jobs but that some operators lacked; thus, in some cases the request may have been threatening. Even the request to participate, absent the specific task being requested, raised uncertainty for operators and called for levels of initiative with which they were unfamiliar and uncomfortable.

These barriers to explication should not be included in the barriers to noticing framework, however. Their influence on noticing is indirect, accomplished only via explication. Given that explication is not necessarily the only or even the best way to influence the barriers and enablers, it makes sense to keep separate these influences on noticing.

The field study revealed several new barriers to noticing. These include socially or organizationally imposed disincentives to search, scripted
alternatives to search, and scripts specifying when to act on the basis of search. Perhaps the field study's most important contribution to the barriers/enablers framework, however, is to reinforce the idea that frames and scripts can come from many sources. Some were idiosyncratic to the individual and situation. Others appeared to be scripted heuristics or frames common to most or all people (e.g., the similarity cue to causality). Still others were supplied and maintained by the organization.

In sum, this dissertation has:

- Defined and drawn attention to noticing and explication, two constructs that have received insufficient attention in past research and whose relationship has not been previously postulated,
- Modeled the inference process, both as a series of activities that people undertake and as a sequential series of states through which variables pass as they are incorporated into mental models,
- Proposed and demonstrated that explication influences the inference process, increasing the likelihood of noticing new variables,
- Proposed a series of phenomena – termed barriers to and enablers of noticing – that influence the inference process and thereby affect the likelihood of noticing,
- Explained how these barriers and enablers have their effect and how explication might influence them,
- Demonstrated that the barriers and enablers mediate and thus partially explain the relationship between explication and noticing,
- Identified a series of variables likely to be important in organizational settings that influence explication and therefore must be considered in designing explication-inducing interventions, and
- Acknowledged that explication is almost certainly not the only, or necessarily even the best, way to influence the barriers and enablers; there may be more effective ways to induce noticing.

6.2. Discussion

This dissertation's findings carry broad implications for managers, methodologists, and theorists, alike. Perhaps more important than the conclusions drawn, however, are the questions raised. In this section, I review the dissertation's most important implications and discuss the key questions it poses.
Chapter 6

Noticing

This research offers a new agenda to managers concerned with their firms' ability to respond to changing environments. It starts by drawing attention to noticing. Its most fundamental contribution to practice is the idea that noticing is a process that can and should be managed. The idea that noticing can be managed raises a series of normative questions of theoretical interest.

I have asserted that most firms suffer from a deficit of noticing. However, we must ask the question, how much noticing is appropriate? The idea that too much or too rapid noticing could distract, destabilize, or prevent complete responses raises the question of whether noticing should be constrained at times. Noticing most likely takes effort and therefore draws attention away from other tasks. How should we weigh the benefits of known tasks, the costs of noticing, and the potential results of noticing? Thus there is a broad normative question, at both the individual and the organizational levels, about when noticing should take place.

A normative model answering this question would have to be contingent on characteristics of both the learner (i.e., the person or organization doing the noticing and responding) and the environment in which it functions. Clearly stability of the environment (relative to the learner's speed of response) is an important consideration. One can imagine that optimal behavior varies considerably (and perhaps discontinuously) at different levels of stability (March, 1991). The learner's ability to handle complexity would also make a difference. Just identifying the contingencies, let alone integrating them into the model, is a major challenge.

The discussion thus far has implied that the noticing process should be managed to ensure that it takes place at times and perhaps to constrain it at others. However, it may be that noticing should always be encouraged and that instead of constraining noticing, managers should focus attention on the how people respond to what they notice. What are the circumstances under which each of these two approaches to managing noticing is likely to be effective?

We also need to ask who should be doing the noticing in organizations. I have asserted that each person represents a firm's best window onto some
aspect of the firm's environment or internal processes. Does this mean everyone should be doing the noticing? Gatekeeper models (e.g., Allen, 1977) suggest that information gathering concentrated in gatekeepers is more effective than widely dispersed responsibility for noticing. These models typically focus on single areas of inquiry and on information sourced from outside the firm; but the decision-makers I know in firms depend heavily on other people for all kinds of information from broad domains, both within and external to the firm. Their experience of the stimuli from which variables can be distinguished is mediated by others, within whom responsibility for noticing must necessarily reside. Those people, too, rely on others. Thus the sources of important information are too diffuse and too unspecifiable for noticing to be left just to those few who take on gatekeeper roles; in a sense, everyone must become a gatekeeper because no one knows where the next important variable will be noticed.

There is at least one fundamental and unresolved methodological problem with noticing as a construct. Measuring it in settings with anything less than the near-complete control of the laboratory is extremely difficult. First, we need to tackle the problem of measuring when people in real-world settings actually notice variables (or how many they notice) without affecting the noticing process; clearly a protocol-based approach could not work. Then we need to make sure that we either: 1) put people in environments that offer equal opportunities for noticing – no small task given that people, because of their different pre-existing mental models and abilities (e.g., color blindness) notice different things – or 2) measure noticing relative to the person's potential for noticing in that environment\(^1\) – another challenging task. The problem of how to measure noticing in any realistic setting has not yet been solved.

**Explication**

Explication's role and the difficulties encountered in the field study raise important questions about how to produce explication in organizations. The

---

\(^1\) Note that even in the laboratory experiment I had to cope with this problem. The two tools used were random assignment and the selection of a capable, homogeneous population from which to draw a sample (see the discussion of the sample in Chapter 3).
learning produced here is only a start. The field study makes clear that we need to consider individual, group, and organization-level influences if we are to design effective explication-inducing interventions and create institutions that maintain explication in the long run. The problem is complex and should be tackled both by theorists and practitioners. The theorist can attempt to deduce, based on existing theories, the practices that might be effective. Practitioners can use their own experience, wits and local knowledge to devise solutions. An interesting approach would be to interest practitioners in explication, set them loose to design and implement their own interventions, and study the process from start to finish.

My work has added a new dimension to the debate on the validity of verbal protocols by focusing attention on the nature and effect – not to mention the practical relevance – of the reactivity noted by scholars such as Nisbett (e.g., Nisbett & Wilson, 1977; Payne, Braunstein & Carroll, 1978; Russo, Johnson, & Stephens, 1989; Smith & Miller, 1978). However, the fact that my work is based on written, rather than spoken, protocols creates distance from this debate that needs to be narrowed. Can spoken protocols have the same effect as the written protocols used in my work? There is some evidence from both the experiment and the field study that the answer is yes. But there is more work to be done before this question is answered.

This dissertation's results should raise healthy skepticism about the validity of research that uses protocol-based measures of process when studying people in real-life problem solving situations. By introducing explication, we as researchers may be fundamentally altering the thought processes of the people we study.

As with noticing, measurement of explication is a challenging problem. Among participants in the field study and members of the experiment's control group, there was so little behavior that resembled explication that it was safe to conclude that the level of explication was zero. But if there had been more, I would have been faced with the question of whether all utterances or scribblings about models constitute explication. This question actually arose at one point and the specific questionable instances were addressed (see "Types of Recording" in Section 3.3.1). Explication involves the jump from theory-in-action to espoused theory. So, clearly, if all a person
does is scan memory for pre-existing models already coded as language and then recite them, he or she is not explicating. We thus are faced with two questions. What behaviors, if any, occur that are not explication but result in the verbalization of what appear to be mental models? Given the answer, how then can we measure explication?

Models of the Inference Process

This dissertation has produced two related models of the inference process. The model of the sequential states through which variables pass during the inference process provides a set of fundamental, low-level constructs from which higher-level ones, such as noticing, can be constructed. This allows for quite precise definitions and should be helpful to theorists who follow.

The process model of inference proposed in Chapter 4 offers opportunities for descriptive and normative work. There is a need to describe the ways that people shift between its three types of search (search within the candidate set, search for new variables, and search for constraining scripts and frames). Furthermore, a normative model needs to be constructed, perhaps integral to the normative model of noticing mentioned above.

The contingencies in these normative models may have the interesting characteristic of being unmeasurable. They may require the learner to make assumptions. Different strategies for making assumptions (e.g., heuristics for deciding when feedback demonstrates a need for improvement, heuristics for deciding when to search within the candidate set versus search for new variables) may thus play an important role in the normative rules that the model yields. Is it possible that we will be able to state normative rules, but only based on variables whose values are unknowable?

Known Barriers to and Enablers of Noticing

The set of barriers and enablers presented here need to be developed further. We need better descriptive models specifying the conditions under which they occur, their rates of incidence, and how to measure them.

Further, frames that act as barriers sometimes enable human information processing because they constrain and thus reduce complexity to manageable
levels. Each barrier likely has an enabling effect under certain circumstances. We need to identify the contingencies that determine when each functions as barrier or enabler and then construct the appropriate contingent normative models to determine their proper role is in the inference process.

Let me illustrate these issues using one specific barrier that I find both intriguing and representative of the broad issues facing this entire area of inquiry. The heuristic that leads people to drop variables from the candidate set when a hypothesis is disconfirmed (and, in general, the issue of discarding variables) has not been addressed in the judgment and decision-making literature. One might offer normative models about dropping variables by appealing to some meta-hypothesis-testing process. However, the participants in the experiment dropped their variables after testing a single hypothesis, which suggests that the true process is sufficiently abbreviated to justify dropping the "meta". Yet, we know that people often experience overconfidence in their judgments (Oskamp, 1982), construct search strategies to help them avoid disconfirming feedback (Klayman & Ha, 1988; Snyder & Swann, 1978; Wason, 1960), and often ignore disconfirming information if they do receive it (Ross et al, 1975). How can we integrate these observations into those made in this experiment? Clearly further descriptive and normative work is needed on this topic.

New Barriers to and Enablers of Noticing

The analysis demonstrates that the barriers to and enablers of noticing can only partially explain the relationship between explication and noticing. This finding motivates further inquiry into the remaining mechanisms by which explication has its effect. My own intuitions suggest that the enablers represent the most fruitful area, in part because they seem underrepresented in this study. An interesting question is why I discovered so few; the answer might be quite revealing and helpful for guiding further inquiry. I have proposed several enablers that might be relevant, including: greater creativity, clarity of thought, use of analogy, creation and consideration of alternate scenarios, or counterfactual reasoning, all mechanisms that did not emerge in this study. More work needs to be done to see if these phenomena do, in fact, function as enablers and to discover still more of them.
Barriers to Explication

There is voluminous work in the area of organizational change management (e.g., Beckhard & Harris, 1987; Beer, 1980; Schein, 1969, 1987) that should help eliminate many of the non-task-specific barriers to explication observed here. When it comes to barriers that arise specifically because the task is explication, theory may have much to offer; but, as I mentioned earlier, my intuitions suggest putting a few practitioners on the problem – thereby taking maximum advantage of their local and in-depth knowledge of their people and organizations – and then studying their actions and results.

Moving Beyond Explication

This dissertation raises two more key questions. The first is about other interventions. I have never claimed that explication is the only or the best way to affect the barriers and enablers identified here. (I am confident that the barriers and enablers are quite fundamental; but this, too, is an open question.) Therefore, I wonder what alternatives to explication, if any, might be more effective, easier to implement, and more easily maintained on an ongoing basis.

Further, the discussion of interventions needs to move to a higher level of analysis. The discussion has frequently introduced group- and organization-level variables. A proper theory of noticing must be multi-level, accounting both for individual-level effects as well as the influence of the groups and organizations in which people work.

6.3. Summary

This dissertation raises many questions; in fact, they are some of its most important outcomes. However, its conclusions also carry far-reaching implications. My hope is that this work draws attention to the importance of noticing, makes salient the possibility of managing it within organizations, and motivates further inquiry among those who care about individual- and organization-level learning and innovation.
Bibliography


Bibliography


Bibliography


Bibliography


Appendix A: The Case Used in the Experiment

Congratulations! You have just been hired by Home Network Fashions (HNF), a small and growing company that offers opportunities for rapid advancement. You joined the company partly because you knew you would be working for Ellen Bolton, a woman who bursts with energy and who has moved up in the company very quickly. She has taken a real liking to you and in private tells you that she's thrilled that the company has finally hired someone with "real management potential." She's determined to have you learn the business from top to bottom, and that's just fine with you.

She tells the story of the company as follows:

"Sandy and Tom Florsheim – you'll meet them, they're around a lot – started HNF in mid-1990. Sandy and Tom had both worked in retail for years. Their next door neighbor was in the cable TV business and one day, it was in May, the three of them were talking over the fence when they had an idea for a new business.

"Essentially it worked as follows: Sandy and Tom searched for manufacturers, distributors, retailers – anybody with overstock that they were prepared to sell at a discount. They focused on fashion accessories like jewelry, watches, and handbags, and some clothing. The items purchased in one week would be advertised on cable TV the next and sold by telephone ordering. Tom did most of the purchasing. Sandy appeared on TV. Within a month they had hired three people to help with handling the telephones, shipping, etc. By 1993 it was a $28 million business.

After they had been running the company for only a year Sandy and Tom recognized that it was taking off and started to create the structure that would allow it to become a large firm. They designed a departmental structure and hired two people who they made vice presidents. In mid 1992, these vice presidents started hiring teams of people to work under them and the firm moved from the Florsheim's farmhouse into an office building and a warehouse nearby.

"That's when I was hired.

"When I got here, there was a lot of chaos around pricing. Basically all people could agree on was that when you raise prices you probably won't sell as much stuff. Meanwhile the company was either carrying all this inventory or selling out by midweek.

"It turns out that pricing for a shopping network is different from any other kind of pricing I've ever done. I just started tracking prices and orders,
estimating a demand curve, and pricing systematically; you'll see that we get about 100 additional orders for every 30 cents I lower price. It made a big difference... and got me promoted. So I've got my approach, but I've got a lot of questions about it. The conflicts about pricing continue to this day. There are still people who disagree with the way I do things and one even quit the company when I was promoted.

"So when it comes to pricing, you'll be able to use my spreadsheets. They'll track prices and orders for you and give you a suggested price – it's based on a simple regression line through past data, but it helps. The spreadsheet also helps me out in calculating advertising costs; but I don't want you to worry about costs yet. You should focus on accuracy.

"To get you started, I'm going to have you work with some past data for a while to get a feel for pricing. I thought that you should start out with setting prices for plastic fashion watches. They're going to be one of the products for which you are responsible."
Appendix B: Instructions for the Experiment

There are two parts to the instructions, one that provides a general orientation that is the first thing that participants read and a second that describes the use of the simulator that they read after the case. Text in square brackets is included in the instructions for the treatment group and excluded for the control group.

General Orientation

Thank you for helping us out with our work. You are participating in an experiment that will help us understand the way people think about everyday business problems involving prediction.

You will be asked to do the following:

I. Answer a few questions about your background, familiarity with supply and demand, and comfort with PCs.

II. Read a short case: The case provides background for a simulation game you will be playing and evaluating.

III. Work with a computer simulation in which you make pricing decisions and where the computer simulates the resulting demand. In each week of the game, you will have a given quantity to try to sell and your goal will be to price such that you sell exactly that amount (you want to avoid both backlogs and inventory).

[As you work, we'd like you to keep a diary of your thoughts.]

IV. Afterwards, we'd like to talk with you briefly about your impressions of the experience, your impressions of the role that the software may have had on your performance, and anything you would like to see changed.

All data that we collect is confidential. This page has a number on it and we use this number to identify you rather than your name. With your permission we would like to record our conversation with you in Step IV; we are the only people who will listen to the tape and, once transcribed, the tape will be erased so that your comments cannot be associated with your voice. All this is really just careful procedure for a task that we hope you'll find interesting and fun.

This set of instructions and the attached case should provide you with all the information you need. If you have trouble with anything, first please refer to the instructions, and then ask David for clarification.
Instructions for the Simulation

You are about to play a pricing simulation game based on Ellen's data. In the game, you will be setting prices for plastic fashion watches. Each week, you will be told how many watches you have to sell and you should set a price in order to sell all of them.

Your goal is to set prices as accurately as possible. If you fail to sell product one week, you'll have to try to sell it in the next. Unlike inventory, orders are not carried over from one week to the next. If you couldn't fill an order one week, you've missed your chance. Pricing too high and pricing too low are penalized equally. Thus you are being evaluated on the basis of how accurately you can match demand each week.

For example, suppose you had 2200 units to sell and you chose a price expecting to sell all of them. Suppose you got only 2000 orders. Then you "missed" by 200 units and that would be your score for the week. You'd also be left with 200 units of inventory to try to sell in the next week in addition to whatever new stock came in. If you got, say 2400 orders, then you "missed" again. You'd sell all your stock and your score for the week would be 200.

After the game, we will discuss your performance in the context of the software you used to help you play it.

Here are more detailed instructions:

1. The simulator should be running already.

2. You should now be looking at a screen with several windows used for displaying information and entering prices. The following describes the information available in each window and how you will enter data to play the game:

**Demand Window:** In the upper right is a graph that shows past data: the prices set plotted against the actual number of orders received. This graph is updated each week after you enter your price. The graph starts with several weeks of data to serve as an initial reference for your decisions.

After you have been playing for a few turns, you'll see points in three different colors on the graph.

- The original points on the graph are red.
- The most recent point is always blue, so you can spot it easily.
- All the other points are green.
Input Window: In the upper left is a window that displays:

- The stock you've got to sell this week (the inventory left over from last week + new stock that your buyers brought in)
- The "suggested price", which is based on a simple linear regression through the points on the demand graph. Once the computer calculates the regression line, it calculates a price to match the stock you have to sell this week.
- There is a place for you to enter the price you want to try.
- What happened last week: The price you set, how many orders you received, the amount of stock you had, and the amount you were actually able to ship.
- Your score for each week is the absolute difference between actual orders and the amount of stock you had. Your cumulated score over the course of the game is also shown.
- Two buttons, one for the advertising cost information (daytime vs. primetime slots, weekend vs. weekday slots) and another to quit the game (but please don't quit from the game; let it reset itself).

Performance Window: In the lower left is a window that shows you how far off you were each week. Unlike your score, this shows whether you had too many or too few orders rather than just magnitude.

3. Enter the price you think will clear your current inventory and then press the Enter key on the numeric keypad. A window will pop up so you can check the price you entered and change it if necessary. If the price is correct, just press Enter again. For the most part, you can play this game just by typing in numbers and pressing Enter.

The system checks to make sure you put in reasonable prices (e.g., numbers, not lower than $20 and not higher than $50) to help avoid typos.

4. Once you enter your price, the software simulates the actual number of orders received and updates the information on the screen.

5. Continue playing the game for 40 weeks at which time the program will automatically reset itself; you won't have to quit.

Please stop now and play the game.
When you are finished (week 40) please let us know.
[Your thoughts are very important to us.
Please help us by maintaining the diary.]
Appendix C: Protocol Instructions

This appendix includes the instructions for the protocol and the first of the protocol's ten pages. The additional nine are identical except for the week numbers:

Please use these pages as a diary for your thoughts as you progress. We ask you specific questions periodically and then give you space to tell us about:

- Your **ideas** about how demand works: your hunches and what you have discovered.
- Your **strategies**.
- **Exclamations** (e.g., whoa!, right on!, @#$%^&*!!!, etc.)
- When you are **surprised**, when things go as expected.
- How **difficult** you think the game is.
- How your **confidence** is changing.
- How you are **feeling**.
- **Anything else** you think we should know about what's going on.
- Etc.

We really want to know what you're thinking about... On the other hand, don't dream up phony stuff just to tell us.
Appendix C

Before entering the price for Week 1, please answer the following:

Overall, my confidence that I can set prices accurately is:

<table>
<thead>
<tr>
<th>Very Low, the barest hunch</th>
<th>Very High, almost 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

With practice, I could improve my current pricing approach:

<table>
<thead>
<tr>
<th>Not a bit</th>
<th>Tremendously</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Things like: What factors are you considering in selecting a price? How do they affect the price you set? How confident are you that each one matters?

Factors; Confidence; Effect on your decisions; Strategies?

Weeks 1-4: New thoughts, hunches, etc.? Do you have a strategy? Etc.

(Go on to the next page before entering a price for Week 5.)
Appendix D: Informed Consent Form

The experiment was approved by the Committee on the Use of Humans as Experiment Subjects (COUHES), as was the following informed consent form:

Thank you for helping me out with this work. You are participating in a research project that will help me understand the way people think about everyday business problems involving prediction.

It is important that you understand that:

- Your participation is entirely voluntary; you may stop participating at any point and you may decline to answer any questions.
- You will be playing a simulation game and discussing it with me afterward.
- I will ask you if I may record our discussion; you are free to ask that this conversation not be recorded.
- Anonymity and confidentiality will be strictly maintained. Your name will never be used in reference to the data nor will you be described in such a way that you might be identified by others. Once the tape is transcribed it will be erased so that no data can be associated with your voice.
- You should be subject to no discomforts or risks as a result of your participation.
- If you have any further questions, you may discuss them with me (David Rabkin), and you may contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, M.I.T. 253-6787.

All this is really just careful procedure for a task that I hope you'll find interesting and fun. Thank you for participating.

- David Rabkin

I have read this form and consent to participate.

Name: __________________________________________ Date: ______________
Appendix E: Field Study Protocol
Instructions

I'm asking you to do two things to help me out over the next few weeks:

1) Keep a notebook in which you take notes on problems that affect the performance of the line or the quality of the product.

2) Review your notebook with me once each week, teaching me about the problems that you have taken notes on and your ideas.

What should go into these notebooks?

I'm interested in what you think about work-related problems or issues. So keep your eyes open and when one comes your way, write down:

1) What day is it?

2) What is the issue? What are its symptoms?

3) Why might it be happening? What are the things that might cause it? Do you think you know what the real cause is?

4) What things might you do about it? What did you actually do? Why?

Please write about things that are happening now, not things that happened in the past. I'm interested in what you think about these things as they are happening.

I'd rather you write up one or two problems completely than write just a little about a lot of different ones. Also, if in our conversations you start telling me about lots of stuff that's not in your notes, I'll ask you to write some more and then talk about it with me.

If it is difficult for you to write in English, write in your own language and we'll translate your notes together when we talk.

How often should you do it and how long should it take?

Please write in the notebook at least once every day, even if it's only for a few minutes.

This should not be a big deal. Try to keep good notes but don't spend a lot of time on it. Probably if you spend 10 minutes each day, you'll do a great job. If you want to write more, that's great, too.

Why use the notebook? Why not just talk?

The notebooks do two things. They can help you think and they help you remember details.
Appendix E

Why should you do this?

One reason is to help me. I'll learn a lot more if you help me. Also, when people do this, they learn new things. Once people get into it, they find it interesting and fun. I hope you will. But remember, you are a volunteer; you don't have to do this.

Will Luther put up with it?

Yes. He's keeping a notebook, too.

Just so that you know:

What you write and say are confidential. The things that you write about and that we talk about are between you and me. I will not show others my notes or identify any of you by name.

Thank you. I really appreciate your help.

- David