ASSESSING THE EFFECTIVENESS OF SYSTEMS-ORIENTED INSTRUCTION FOR PREPARING STUDENTS TO UNDERSTAND COMPLEXITY

By

RICHARD RANDALL PLATE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2006

To Muriel Plate

ACKNOWLEDGMENTS

It has been a pleasure to work with and for my advisor, Leslie Thiele. His keen insights and subtle guidance added much to my experience. In addition, his confidence in my ability to apply my ideas regarding systems-based instruction played an integral part in the acquisition of the data for the University of Florida study. I thank the members of my supervisory committee: Mark Brown, Stephen Humphrey, Anna Peterson, and especially Martha Monroe, whose open-door policy caused her to lose a lot of hours conjecturing with me about this study. I am also thankful to Frank Day. The chance he took on me years ago marks a key moment in my development—both intellectually and in general.

I would also like to aknowledge the cooperation of Portland Public Schools and the teachers who were involved with this study. In particular, I am indebted to Tim Taber and Mary Scheetz who were a great help in my efforts to set up data collection for the Portland study. It is difficult (and painful) to imagine how that portion of the research would have gone without their assistance.

And of course, I would like to acknowledge the support of family and friends who were able to make even the most stressful times worth smiling about. Lastly, I should thank Lily, my constant companion during the writing of this dissertation. Any insights contained herein that are worth mentioning probably came after a walk with her.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS 4

LIST OF TABLES 9

LIST OF FIGURES 11

ABSTRACT 13

CHAPTER

1 INTRODUCTION 15

1.1 A Riddle 15
1.2 Outline of This Research 20
Notes 22

2 SYSTEMS THINKING AND CURRICULUM DEVELOPMENT 23

2.1 Introduction 23
2.2 A Short History of Crises in Education 26
2.3 The Co-Adaptive Curriculum 33
2.4 A New Crisis? 40
2.5 Conclusion 44
Notes 45

3 SYSTEMS-ORIENTED INSTRUCTION 49

3.1 Beyond Reductionism 49
3.2 A Systems Perspective 52
3.3 Nonlinear Causality 53
3.4 Causal Loops 56
3.5 Identifying Positive and Negative Feedback in Systems Diagrams 59
3.6 Systems Archetypes 60
3.7 Systems Thinking and Transfer of Learning 64
3.8 Why Is Learning Systems Concepts Important? 68
3.9 Conclusion 71
Notes 71

4 SYSTEMS THINKING IN THE ROLE OF ENVIRONMENTAL DECISION-MAKING 83

4.1 Introduction 83
4.2 Simple Heuristics 86
4.3 Framing Problems: The Use of Mental Models 90

4.4 Decisions in the Context of Complexity 98
4.5 Conclusion 105
Notes 107

5 COGNITIVE MAPPING AS A METHODOLOGY FOR ASSESSING THE EFFECTIVENESS OF SYSTEMS-ORIENTED INSTRUCTION 110

5.1 Introduction 110
5.2 Cognitive Mapping 113
5.3 Contemporary Cognitive Mapping 114
5.4 Cognitive Mapping as a Tool for Environmental Education 115
5.5 Concept Maps 118
5.6 Cause Maps 120
5.7 Conceptual Content Cognitive Maps 123
5.8 Interview Methods 124
5.8.1 Developing a Scenario 125
5.8.2 Creating Concept Cards 127
5.8.3 3CM Exercise 129
5.8.4 Causal Mapping 130
5.9 Conclusion 132
Notes 134

6 EMPIRICAL ASSESSMENT OF SYSTEMS-ORIENTED INSTRUCTION 138

6.1 Introduction 138
6.2 Description of Calculations 138
6.2.1 How to Read this Chapter 138
6.2.2 Analysis of 3CM Maps 139
6.2.2.1 Was there a difference in the average number of concepts chosen by each group? 139
6.2.2.2 Was there a difference between groups regarding which concepts were chosen for the map? 139
6.2.2.3 Was there difference in how the participants from each group categorized their chosen concepts? 140
6.2.3 Analysis of Cause Maps 140
6.2.3.1 Did the participants include web-like causality in their causal maps? 142
6.2.3.2 Did the participants include causal loops in their causal maps? 144
6.2.3.3 Were there differences in the specific connections identified by participants from different groups? 145
6.2.3.4 How successful were the participants in accurately representing the issue described in the article? 146
6.3 Study 1: Undergraduates from the University of Florida 148
6.3.1 Study Participants 148
6.3.2 Analysis of 3CM Maps 149
6.3.2.1 Was there a difference in the average number of concepts chosen by each group? 149

6.3.2.2 Was there a difference between groups regarding which concepts were chosen for the map? 149
6.3.2.3 Was there a difference in how the participants from each group categorized their chosen concepts? 150
6.3.3 Analysis of Causal Maps 151
6.3.3.1 Did the participants include web-like causality in their causal maps? 151
6.3.3.2 Did the participants include causal loops in their causal maps? 151
6.3.3.3 Were there differences in the specific connections identified by participants from different groups? 152
6.3.3.4 How successful were the participants in accurately representing the issue described in the article? 153
6.3.4 Summary and Discussion 154
6.3.4.1 Relative Focus of Concepts and Connections 154
6.3.4.2 Structural Differences and Comparison to Expert Maps 155
6.3.4.3 Student Opinions 155
6.3.4.4 Shortcomings of the Study 157
6.4 Study 2: Portland Study 158
6.4.1 Description of Participants 158
6.4.1.1 Comparison of study groups 158
6.4.1.2 Systems curriculum at School A 160
6.4.2 Analysis of 3CM Maps 162
6.4.2.1 Was there a difference in the average number of concepts chosen by each group? 162
6.4.2.2 Was there a difference between groups regarding which concepts were chosen for the map? 162
6.4.2.3 Was there difference in how the participants from each group categorized their chosen concepts? 164
6.4.3 Analysis of Causal Maps 165
6.4.3.1 Did the participants include web-like causality in their causal maps? 165
6.4.3.2 Did the participants include causal loops in their causal maps? 166
6.4.3.3 Were there differences in the specific connections identified by participants from different groups? 166
6.4.3.4 How successful were the participants in accurately representing the issue described in the article? 170
6.4.4 Summary and Discussion 173
6.4.4.1 Relative Focus of Concepts and Connections 173
6.4.4.2 The Role of Public Perception 175
6.4.4.3 Structural Differences between Maps 175
6.4.4.4 Linear Thinking and Unlikely Connections 176
6.4.4.5 Comparison to Experts 177
6.5 Conclusions 177
6.5.1 Lessons Learned Regarding Calculations 177
6.5.2 Group Comparisons 178
Notes 179

7 EPILOGUE 199

7.1 Dohaku's Story 199
7.2 Recommendations..200

ARTICLE USED FOR MAPPING EXERCISE 204

SCRIPT USED FOR MAPPING EXERCISE..207

Part 1:..207
Part 2..209

LIST OF REFERENCES..216

BIOGRAPHICAL SKETCH 229

LIST OF TABLES

Table page

2-1 Jobs are changing due to shifts in organization and management....................................47

6-1 Concept cards listed according to themes....................................181

6-2 Rules for scoring participant maps against expert agreement....................................182

6-3 Concepts chosen during pre- and post-tests....................................182

6-4 Stable categories chosen by participants....................................183

6-5 Percentage of participants identifying causal links in 1^{st}-order matrices among undergraduates....................................184

6-6 Percentage of participants identifying causal links in 5^{th}-order matrices among undergraduates....................................184

6-7 Comparisons of total similarity scores for pre- and post-tests....................................185

6-8 Total similarity scores between expert maps....................................185

6-9 Comparisons of rubric scores for pre- and post-tests....................................185

6-10 Survey responses regarding systems concepts....................................186

6-11 Basic statistics of schools included in the study....................................186

6-12 Concepts included in cognitive maps by each group....................................187

6-13 Differences in the percentages at which concepts were chosen by eighth-grade groups...188

6-14 Percentage of students from each group that chose concepts for particular themes..........188

6-15 Hierarchical clustering results for seventh-grade participants....................................189

6-16 Hierarchical clustering results for eighth-grade participants....................................190

6-17 Web-like causality mean scores....................................190

6-18 Percentage of participants identifying causal links in 1^{st}-order matrices among seventh graders....................................191

6-19 Percentage of participants identifying causal links in 5^{th}-order matrices among seventh graders....................................191

6-20 Differences in theme connections among seventh graders....................................191

6-21 Percentage of participants identifying causal links in 1st-order matrices among eighth-graders 192

6-22 Percentage of participants identifying causal links in 5th-order matrices among eighth-graders 193

6-23 Differences in theme connections among eighth graders 194

6-24 Comparison of Seventh-Grade Total Similarity scores* 194

6-25 Comparison of eighth-grade total similarity scores 194

6-26 Comparison of seventh-grade rubric scores 195

6-27 Comparison of eighth-grade rubric scores 195

6-28 A8 results with increasing variability at higher-orders 195

6.29 Scores from two members of group A8 195

6.30 Summary of results from both studies and experts 196

LIST OF FIGURES

Figure page

2-1 Four aspects of meeting the educational needs of society 48

2-2 Oscillations in the Lotka-Volterra predator-prey model 48

3-1 Two models for causality—causal chain versus a causal web 73

3-2 Gore's causal connections 74

3-3 Cold war perspectives of the arms race 75

3-4 Household thermostat as an example of a negative (balancing) feedback loop 76

3-5 Cold war arms race expressed with arrows identified 76

3-6 Diagram of a system with one prey species and one predator species 77

3-7 Generalized diagrams for reinforcing (left) and balancing (right) loops 77

3-8 Escalation archetype 78

3-9 Shifting the burden 79

3-10 Limits to growth 80

3-11 Tragedy of the commons 81

3-12 Exponential population growth 82

4-1 Simple system (Forrester 1996) 108

4-2 Complex system (Forrester 1996) 108

4-3 Change in demand of beer of the thirty-six simulated weeks of the Beer Distribution Game. (Sterman 1989) 109

4-4 Orders recorded from four different teams in the Beer Distribution Game. Each tick mark on the y-axis represents ten cases. From top to bottom, the graphs represent orders from the retailer, the wholesaler, the distributor, and the factory. (Sterman 1989) 109

5-1 A concept map for water taken from Novak (1984) 135

5-2 A cause map depicting interactions in an interview (Eden et al, 1992). Note that "…" should be read as "rather than." 136

5-3 A 3CM map representing a non-EE teacher's view of environmental education (Kearney 1999) 137

6-1 Types of web-like causality....................197

6-2 Sample causal maps with equivalent link densities....................197

7-1 Positive Feedback Loop between Systems-oriented Instruction and Assessment 203

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

ASSESSING THE EFFECTIVENESS OF SYSTEMS-ORIENTED INSTRUCTION FOR PREPARING STUDENTS TO UNDERSTAND COMPLEXITY

By

Richard Randall Plate

December 2006

Chair: Leslie Thiele
Major Department: Interdisciplinary Ecology

Complex systems abound. They affect every aspect of our lives, from our social and economic endeavors down to the food we eat and the air we breathe. And for the first time in history, the effects of human activity on large-scale environmental systems is occurring rapidly enough to be observed over only a few decades. In addition, the inability of individuals to apprehend even fundamental aspects of these systems has been well documented.

This research presents systems-oriented instruction as a promising pedagogical tool for preparing students to understand complex social and ecological systems. A methodology is presented using cognitive mapping to evaluate how systems-oriented instruction affects the way students learn about complex systems. This methodology is used in two studies, one with undergraduate students from the University of Florida (UF) and a second with middle school students from Portland, Oregon. In each study, participants read an article about a hypothetical fishing controversy involving the interaction of social, economic, and ecological processes and worked through a cognitive mapping exercise during which they were able to express their interpretation of the situation described in the article. The cognitive maps produced by the students were then subjected to a battery of quantitative and qualitative evaluations.

The UF study involved a pre-test/post-test format, in which the students were evaluated at the beginning and end of a semester class using systems-oriented instruction. In the Portland study the cognitive maps of students who have been receiving systems-oriented instruction were compared to those of students who have been receiving conventional instruction. In the UF study, results showed that students' ability to apprehend key aspects of the situation described in the article had improved significantly over the semester. In the Portland study the systems groups displayed a greater understanding of the situation described in the article than the control groups. While the differences observed are not definitive enough to make strong claims about systems-oriented instruction based only on these studies, they are strong enough to warrant further studies assessing systems-oriented instruction's worth as a pedagogical tool. The methodology described here is presented as a model for those studies.

CHAPTER 1
INTRODUCTION

1.1 A Riddle

What does World War II anti-aircraft artillery have to do with the essence of life? If you answered, "Everything, to anyone flying an aircraft during World War II," you are only partly right. The rest of the answer involves one of the more promising and least pursued developments in the context of education for sustainability. Such an answer requires some explanation. A lot of explanation. Perhaps a dissertation's worth.

Let us begin our story at the Aberdeen Proving Ground during the First World War, where young mathematician Norbert Wiener is working alongside other mathematicians to compute firing tables for the United States Army's artillerymen (relating gun elevation to factors such as type of missile and distance to target). After WWI, Wiener took a position at MIT, where he made a name for himself as a distinguished mathematician before returning to military service in WWII—this time working on anti-aircraft technology, which presented a far more challenging problem. The speed of the newest models of airplanes made them extremely difficult targets for ground-to-air defense. Wiener's solution to the problem involved a feedback loop between the radar and the radar-guided artillery.

This solution made Wiener curious about the role that feedback loops play in other functions, including cognitive, biological, and social processes. The result of this curiosity was a series of conferences (the Macy Conferences), beginning in 1942 and involving a number of notable scholars including psychiatrist Warren McCulloch, anthropologists Margaret Mead and Gregory Bateson, electrical engineer Julian Bigelow, and mathematician John von Neumann. During these conferences, the group explored the role of feedback in a broad range of systems, with Wiener concluding that biological and social systems rely on feedback in a way analogous to his radar-guided guns. All of these systems, he explained, “have sensory receptors as one stage in their cycle of operation.” In these systems, the “performed action on the outer world, and not merely their intended action, is reported back to the central regulatory apparatus. This complex behavior is ignored by the average man, and in particular does not play the role it should in our habitual analysis of society” (1954, p. 27). This thesis on the importance of feedback loops became the basis for the new field of cybernetics, the science of goal-directed systems.

During this same period, philosopher/biologist Ludwig von Bertalanffy was busy with a less violent conflict—that between the mechanists and the vitalists. This conflict was over the definition of life: what separates non-living matter from living matter. The mechanist response to this question was that the things we refer to as life and consciousness were simply epiphenomena resulting from physical and chemical processes. The vitalists, on the other hand, believed in a life force that animated all living matter. Like others in the debate, Ludwig von Bertalanffy did not believe that chemical and physical laws alone could explain the behavior of living matter, but was unconvinced by the idea of an unperceivable force as an explanation. A more satisfying view for von Bertalanffy required changing the way one looked at the processes of life. He explained:

> There is no “living substance” in the sense that lead, water, or cellulose are substances, where any arbitrarily taken part shows the same properties as the rest. . . . So long as we consider the individual chemical reactions that take place in a living organism we are unable to indicate any basic difference between them and those that go in inanimate things. . . . But a fundamental contrast is found when we consider not single processes, but their totality within an organism or a partial system of it, such as a cell or organ. Then we find that all parts and processes are so ordered that they guarantee the maintenance, construction, restitution, and reproduction of the organic systems. (Bertalanffy 1949, p. 13)

Thus for von Bertalanffy, the vitalist debate arose because both sides had attempted to define life solely at the molecular level without taking into account the organization and interrelationships of those molecules within living systems.

Together, Norbert Wiener and Ludwig von Bertalanffy represent the foundations of systems thinking, not because they were the only ones to think of these ideas,[1] but because of their integral roles in developing a formal set of concepts that could be applied in a number of different areas. These ideas have gone far beyond cybernetics and theoretical biology, having been applied to economics (Boulding 1953; Daly 1973), cognitive science and epistemology (Maturana and Varela 1987), and educational administration (Banathy 1991). Particularly notable in the present context are the efforts of H.T. Odum and Jay Forrester.

H.T. Odum, coming from a systems ecology background inspired by Alfred Lotka (1945), applied systems concepts to the study of ecosystems and developed an energy circuit language "combining kinetics, energetics, and economics" (Odum 1983, p. ix). Jay Forrester founded the field of system dynamics, a methodology for studying complex systems, and applied it to business management and urban development, among other topics (1961; 1969). These two men distinguished themselves not only by establishing formidable reputations in their respective fields, but also by their focus on applying systems concepts to curriculum development.

Odum and Forrester believed, independently, that systems concepts, because of their broad applicability, could become useful pedagogical tools. Odum explains, "If the bewildering complexity of human knowledge developed in the twentieth century is to be retained and well used, unifying concepts are needed to consolidate the understanding of systems of many kinds and to simplify the teaching of general principles" (Odum 1983, p. ix). Forrester expands on this idea, calling for the use of systems dynamics as "as a foundation underlying education in kindergarten through 12th grade" (1997, p. 4). Others have echoed this call for orienting curriculum toward systems concepts. For example, Peter Senge (1990) presents systems thinking as a fundamental part of education for employees working in successful 21st-century corporations, and C.A. Bowers (1995) cites systems thinking as a foundation for creating an ecologically sustainable culture.

And Bowers is not alone. Systems thinking has often been cited as a means for addressing our most pressing environmental issues. John Sterman and Linda Booth Sweeney (2002) document the role that systems thinking plays in understanding global climate change, arguably the most pressing environmental concern today. David Holmgren characterizes his system of sustainable agriculture (called permaculture) as "the use of systems thinking and design principles that provide the organising framework for implementing the above vision [of providing] for our needs, while increasing the natural capital of future generations" (2004, p. 1). In the context of reducing water pollution, Samuel Moore and Larry Ausley cite systems thinking as "a particularly effective approach to solving aquatic toxicity problems [associated with industrial manufacturing] because there was very little data to indicate a linear relationship between removal of traditional pollutants and reduction in aquatic toxicity" (2004, p. 590). And in the context of water management, Mark Everard explains, "By application of systems thinking, sustainable solutions to eutrophication of still waters may be achieved not by local 'end of pipe' technological responses to perceived local problems, at net cost and with their own environmental impacts, but by basing wise decisions upon the dynamics of the wider water cycle" (1999, pp. 33-34).

Also, a growing number of scientists and educators include systems thinking as a key aspect of a curriculum that addresses the complexity of today's world. The National Science

Foundation (NSF), which lists as one of its four primary functions, supporting “science and engineering education programs at all levels” (NSF 2001), cites systems thinking not only as a goal, but as a guiding principle. In the National Science Foundation’s five-year strategic plan for FY 2001-2006, the authors describe the NSF as “embracing change through effective systems-thinking and appropriate feedback mechanisms” (NSF 2001). Similarly, the multi-million dollar educational program GLOBE—jointly funded by the NSF and NASA—lists three core missions: “to improve science education, to increase environmental awareness, and to contribute to understanding of the Earth as a system” (GLOBE 2005).

In addition to funding agencies, a number of state governments have included systems thinking in their standards for public education. For example, among Colorado’s science standards for high school is “identifying and describing the dynamics of natural systems (for example, weather systems, ecological systems, body systems, systems at dynamic equilibrium)” (Colorado Department of Education 1995). New York’s science standards include one on interconnectedness: “Through systems thinking, people can recognize the commonalities that exist among all systems and how parts of a system interrelate and combine to perform specific functions” (New York State Education Department (1999). And Florida includes among its social studies standards the knowledge of “how humans overcome ‘limits to growth’ imposed by physical systems” (Florida Department of Education 2005). As we shall see, the mention of limits to growth is a direct allusion to a famous systems concept.

And finally, systems thinking has received international attention in the context of education for sustainability. In December 2002, the United Nations adopted resolution 57/254 on the United Nations Decade of Education for Sustainable Development (2005-2014), designating UNESCO as the lead agency for “providing recommendations for Governments on how to promote and improve the integration of education for sustainable development" (United Nations 2003). Commenting on the resolution, Koïchiro Matsuura, Director-General of UNESCO, described a "new vision of education" that "emphasizes a holistic, interdisciplinary approach to developing the knowledge and skills needed for a sustainable future" (United Nations 2002). And speaking on the topic of sustainability, former Secretary General for the United Nations, Maurice Strong suggests that systems-oriented instruction can help students "to make sense of the complexity of human-environment interactions, enabling them to make connections not previously noticed in many areas of their lives" (2001, p. xiii).

But what is the outcome of this attention. For all the buzz-word value that terms like *holistic* and *interconnectedness* hold in the vernacular of proponents for ecological literacy, the majority of attempts to actually apply systems-oriented instruction appear in the context of business management. Only a handful of schools across the country can claim to be implementing systems-oriented instruction at an institutional level. If systems concepts are as useful as many claim for understanding the complexity of our society in general and our environmental challenges in particular, then surely they should be included in curriculum more commonly than they currently are. But presently, we know little about exactly how useful these concepts are in the context of education for sustainability, nor do we know how useful they can be or how they might be implemented to maximize their utility.

1.2 Outline of This Research

With this research, I provide the reasoning behind much of the attention garnered by systems concepts, and I present a methodology for assessing the usefulness of systems-oriented instruction in the context of understanding the complex environmental challenges that we currently face. In Chapter 2, I place contemporary calls for including systems concepts in curricula within the context of curriculum development in the United States. I argue that far from being a radical curricular shift, the implementation of systems-oriented instruction represents a logical next step in the context of curriculum development.

In Chapter 3, I provide a description of what I mean by systems-oriented instruction. This chapter includes an introduction to basic systems concepts and a description of how those concepts can be implemented into the curriculum. Again, we see that systems-oriented instruction does not imply a radical departure from conventional curricula. The course subjects and content need not change very much at all. What does change is the approach to the content and the connections drawn between the subjects.

In Chapter 4, I provide an idea of how systems-oriented instruction might change student performance on a practical level. Prior research regarding how individuals interpret information about complex systems is presented in the context of decision-making. Here, we shall see that while there are few studies assessing the effectiveness of systems-oriented instruction in improving students' ability to understand complexity, there is no shortage of studies showing that individuals, regardless of education level, generally perform poorly in attempts to understand examples of complex systems.

Recognizing that there is room for improvement in this area, I address in the next two chapters the question of whether or not systems-oriented instruction can produce that improvement. Chapter 5 provides a review of cognitive mapping techniques and a description of the specific mapping methodology used in the empirical portion of this research. This methodology is presented as a reasonably efficient procedure for producing quantitative data regarding how students comprehend new information about complex systems. In Chapter 6, I describe numerous qualitative and quantitative techniques for evaluating the cognitive mapping data. The results of two separate studies are reported, and the significance of those studies is discussed.

Notes

[1] Max Wertheimer and Kurt Koffka developed in the context of Gestalt psychology a response to the reductionist/vitalist debate similar to that of von Bertalanffy. See, for example, Koffka's (1935) *Principles of Gestalt Psychology*.

CHAPTER 2
SYSTEMS THINKING AND CURRICULUM DEVELOPMENT

> One can hardly believe there has been a revolution in all history so rapid, so extensive, so complete. Through it the face of the earth is making over, even as to its physical forms; political boundaries are wiped out and moved about, as if they were indeed only lines on a paper map; population is hurriedly gathered into cities from the ends of the earth; habits of living are altered with startling abruptness and thoroughness; the search for the truths of nature is infinitely stimulated and facilitated, and their application to life made not only practicable, but commercially necessary.
>
> -John Dewey 1902

2.1 Introduction

In his 2006 State of the Union address, George W. Bush announced his new educational program, the American Competitiveness Initiative, committing over $136 billion over the next ten years "to encourage innovation throughout our economy" (Bush 2006). A United States Department of Education (D.O.E.) report, titled *Answering the Challenge of a Changing World: Strengthening Education for the 21st Century*, explains the administration's view of "innovation":

> To Americans, innovation means much more than the latest gadget. It means creating a more productive, prosperous, mobile and healthy society. Innovation fuels our way of life and improves our quality of life. And its wellspring is education. (U.S. D.O.E. 2006, p.3)

Educator David W. Orr has sharply criticized Americans' view of innovation, what he calls "technological fundamentalism"—a "kind of technological immune deficiency syndrome that renders us vulnerable to whatever can be done and too weak to question what is that we should do" (2002, p. 63). Nonetheless, Orr would likely find much to agree with in the above characterization, for he aims his criticisms specifically at Americans' focus on the latest gadget and the general disregard for a more healthy society. In Orr's words, Americans favor "innovations that produce fast wealth, whatever their ecological or human effects . . . on long-term prosperity" and neglect innovations "having to do with human survival" (2002, p. 69).

Both Bush and Orr cite the need for education to respond to the changing global situation. The difference is a matter of focus. For the current Bush administration, a more productive, prosperous, and healthy society implies the need to graduate students who are prepared to compete in a global market and maintain a high level of national security. While Orr acknowledges the importance of enabling individuals "to compete more favorably in the global economy," he suggests that there are "better reasons to rethink education" (1994, p. 26)—among them, the challenges of stabilizing world population, reducing greenhouse gas emissions, protecting biodiversity, and managing renewable resources sustainably. In short, Orr explains, students today "must begin the great work of repairing as much as possible, the damage done to the earth in the past 200 years of industrialization" (1994, p. 26).

For Bush, a renewed focus in science and math represents the best way to meet current educational challenges; for Orr, the most important curriculum change involves using environmental lessons to integrate school subjects, turning them into a cohesive whole and producing ecologically literate graduates. With these suggestions, both Bush and Orr echo a charge made often during the hundred-year history of curriculum development in the United States: our schools do not adequately prepare their students to meet the demands of contemporary society. Reasons cited for this failure can be put into two broad categories. One, a critic may find the educational theory to be lacking. That is, new research on education or cognitive development may have produced findings that point toward new developments in curriculum. And two, the educational context—that is, the demands placed upon students by society—may have changed, necessitating a corresponding change in curriculum. These categories are not mutually exclusive. In practice critics of curriculum often cite some combination of the two.

The focus of this research favors the second category. While I point to some practical advantages reported by educators regarding student learning in the context of systems-oriented instruction and to a handful of studies on human perception, I do not attempt to introduce a new learning theory. My argument is not that conventional curricula used at primary, secondary, and tertiary levels of education are fundamentally wrong. Rather, I believe, along with Bush and Orr, that conventional curriculum does not prepare students for the new challenges they will face as a result of industrialization and, more recently, globalization. In this chapter I review how the needs of both the individual and the society have shifted over the last century and trace the response to these shifts in curriculum development.

A brief note on terminology is in order. The field of curriculum development has come to mean something significantly different from its original meaning. In the seminal text *The Curriculum* Franklin Bobbitt offers two definitions for curriculum: "(1) it is the entire range of experience, both undirected and directed, concerned in unfolding the abilities of the individual; or (2) it is the series of consciously directed training experiences that the schools use for completing and perfecting the unfoldment." Bobbitt explains, "Our profession uses the term usually in the latter sense" (1918, p. 43; quoted in Jackson 1992, p. 7). More recently, the field of curriculum development (or simply curriculum) has been broadened to include investigations into the former definition, focusing on the "hidden curriculum" (i.e., implicit lessons that students receive during their educational experience). Thus contemporary studies in the curriculum often analyze the school experience in terms of gender, race, or politics, just to name a few. While this is a rich area of research, I will confine the present study to the latter definition, staying, as Phillip Jackson describes "within the single tradition of curriculum specialist as advice giver to practitioners" (1992, p. 27).

Efforts regarding this more narrow view of curriculum development have been criticized for lacking historical perspective (Bellack 1969; Davis 1976; Moore et al. 1997). Veteran educators, having over the course of their careers seen countless educational fads move in and out of fashion, have perhaps earned the skepticism with which they often view new educational techniques and tools. As one contemporary educational theorist admits in the introduction of his text on a new theory, "The promise of a new educational theory…has the magnetism of a newspaper headline like 'Small Earthquake in Chile: Few Hurt'" (Egan 1997, p. 2). Fortunately, I am not trying to introduce a new educational theory here. My objective is much more humble: to point out a specific shortcoming common in contemporary curricula in our public schools and universities and to suggest systems-oriented instruction a promising tool for correcting that shortcoming.

In the following section, I provide a general history of curriculum development since the end of the nineteenth century. From this context, Section III emphasizes specific aspects of curriculum change and describes its relationship with social change. In Section IV I describe contemporary social challenges and discuss how systems-oriented instruction might address these challenges.

2.2 A Short History of Crises in Education

In The Saber-Toothed Curriculum *Harold Benjamin (1939) tells the story of a Paleolithic educational system. In this particular tribe schools focused on three skills that were crucial for their young people to learn: fish-grabbing, horse-clubbing, and tiger-scaring. The first skill provided food, the second provided both food and skins, and the third was a matter of safety. Under this system, the young people were taught the skills they needed to prosper in the future and to help the tribe prosper in the future.*

There is no shortage of reported crises in the history of curriculum design in this country. The first came at the turn of the twentieth century. For most of the nineteenth century American educators emphasized the traditional Latin and Greek curricula of the classics. The field of faculty psychology (also called mental discipline) provided the scientific foundation for these traditional curricula. Proponents of faculty psychology viewed the mind as a muscle to be exercised by memorization and recitation (Pinar et al. 1995, 71-73). Often cited by curriculum history scholars, *The Yale Report on the Defense of the Classics* expresses the motivation behind much of the traditional curriculum: "Familiarity with the Greek and Roman writers is especially adapted to form the taste, and to discipline the mind, both in thought and diction…. It must be obvious even to the most cursory observer, that the classics afford materials to exercise talent of every degree" (Yale Report 1828, pp. 35-36).

Nineteenth century critics voiced their objection to the traditional curriculum on two fronts. First, the choice of curriculum was cited as evidence of over-emphasis of public high schools on college preparation. At its outset American public education was designed to give all of its citizens an equal opportunity to education. The traditional curriculum, critics charged, failed to address the needs of students not bound for college—a group that in the 1889-1890 school year comprised 85% of American high school students (Meyer 1967, p. 405; Tanner and Tanner 1990, p. 68). This problem worsened as enrollment in city schools skyrocketed due to large waves of immigrants and the general trend toward urbanization (Cremin 1961, p. 20; Ornstein and Levin 2000, p. 152). School was no longer just for the elite; it was for the masses. As such, it was subject to criticisms of practicability.

Secondly, and perhaps more importantly for contemporary discussion, the focus on memorization and recitation was seen as responsible for shortcomings in students' basic reasoning skills. As a result, Americans were seen as particularly susceptible to persuasive rhetoric. Charles Eliot, the president of Harvard at the turn of the century and a leading curriculum scholar, argued that the traditional curriculum would not "protect a man or woman…from succumbing to the first plausible deduction or sophism he or she may encounter." He continued, "One is fortified against the acceptance of unreasonable propositions only by skill in determining facts through observation and experience, by practice in composing facts or groups of facts, and by the unvarying habit of questioning and verifying allegations and of distinguishing between facts and inferences from facts, and between a true cause and an antecedent event" (Eliot 1892, pp. 75-76). Other leading scholars of the period expressed fear that most Americans were being educated by the "cheap newspapers" that were shaping public opinion "via the emotional appeal of sensational events" (Tanner and Tanner 1990, p. 91)[1].

By the turn of the century the work of Edward L. Thorndike, who is credited with the rise of experimental psychology in education, had discredited faculty psychology, taking away the traditionalists' scientific foundation (Cremin 1961, Pinar et al. 1995). However, Tanner and Tanner (1990) suggest that a more powerful force of change was at work as well. They explain that the traditional curriculum, as well as the faculty psychology, "had originally evolved to serve an aristocratic society and, in addition to being absolutely unfounded from a scientific standpoint, it did not meet the new social and industrial demands of a democratic society. These demands, rather than the findings of experimental psychology, proved to be the most powerful argument against mental discipline" (1990, p. 110). As a result of this failure, in 1900 the vast majority of students (almost 90%) enrolled in public schools dropped out before graduating high school, citing a lack of need for what was being taught (Tanner and Tanner 1990, p. 72).[2]

The result of the criticism and public dissatisfaction was a broadening of the curriculum and a shift of emphasis toward the learner. Classical studies gave way to contemporary and vocational studies, and it was in this context, during the first few decades of the twentieth century, that John Dewey's progressive ideas of education came to exert more influence on curriculum. The focus shifted from classical materials and languages to meeting the individualized needs of the student. This shift is evidenced in the catch phrases like "the needs of learners," "teaching children, not subjects," and "adjusting the school to the child" (Cremin 1961, p. 328).

In addition to bringing the learner into focus, this period also saw a new emphasis on the connection between education and the welfare of society as a whole. More specifically, Lester Frank Ward—and later Dewey—proclaimed the potential for a school system to create social change (Pinar et al. 1995, p. 104). Many educational reformists believed that the industrialization of society not only created a demand for a skilled public, but also "had dissolved the fabric of community leaving alienation in its wake" (Cremin 1961, p. 60). In this context the role of school broadened to include preparing students not only for a career, but also for understanding their career in the context of the larger social system.[3] A statement from the oft-cited 1917 National Education Association (NEA) report, "The Cardinal Principles of Secondary Education," illustrates how the inclusion of both the individual student and the larger connections to society had become part of the institutionalized focus of education: "[E]ducation in a democracy…should develop in each individual the knowledge, interests, ideals, habits, and powers whereby he will find his place and use that place to shape both himself and society toward ever nobler ends" (1918, p. 157)

This shift of focus met with its own set of critics who believed that the progressive influence was making education soft, ultimately depriving students of a sound background in basic lessons. Such criticism came to a head after the end of World War II, energized by new concerns regarding communist expansion and the rise of the Soviet Union. These concerns combined with budgeting problems brought about by the war, rampant inflation, and ever increasing industrial demands for a trained, intelligent workforce to usher in "the deepest educational crisis in the nation's history" (Cremin 1961, p. 339). Numerous texts, including Bernard Iddings Bell's *Crisis in Education* (1949) and Arthur Bestor's *Educational Wastelands* (1953), accused education of misplaced emphasis on social and emotional matters to the detriment of fundamental, academic skills. Curriculum scholar Hilda Taba suggested, "Public education today is facing a crisis which may be deeper and more fundamental than any preceding one" (1962, p. 1).

The launch of the Soviet satellite Sputnik in 1957 galvanized concerns regarding Soviet supremacy over the United States, reinforcing the back-to-the-basics attitude in school and placing science education at the center of concern. Congress responded with the National Defense Education Act (NDEA) in 1958, directing federal funding to improve science and mathematics curricula and increase opportunities for exceptional students seeking training in critical scientific fields.[4] The National Academy of Sciences organized a conference at Woods Hole in Cape Cod, gathering together leading psychologists, scientists, and mathematicians to discuss how best to help American students become the scientific leaders of the future. Jerome Bruner, chair of the conference, published the results in the physically unassuming book, *The Process of Education* (1953).

Weighing in at ninety-two pages, *The Process of Education* is widely viewed as the most influential curriculum text of its time (Tanner and Tanner 1980, Willis et al. 1993, Pinar et al. 1995, Marshal et al. 2000), during a period referred to as "one of the largest and most sustained educational reform movements in American history" (Silberman 1970, p. 158). Bruner devoted the opening two chapters to the importance of teaching the "structure of the disciplines," meaning basic concepts regarding each subject (e.g. chemistry, language, mathematics). This phrase—and indeed the whole text—was understood by back-to-the-basics proponents as supporting the need for a rigidly defined discipline-centered curriculum. Richard Hofstadter (1961) even stretched Bruner's ideas to support Hofstadter's own arguments for a resurgence of faculty psychology. While Bruner's ideas are at times overly rigid,[5] Marshal et al. point out, "To be fair, Bruner's ideas were far more complex than the manner in which they were eventually employed" (2000, p. 57). We will revisit the complexity of Bruner's ideas later in the chapter.

For all of its influence, Bruner's strict, top-down system of curriculum development did not match the 1960s trend toward liberation. By 1969, educators had come to associate the top-down curriculum with "the military-industrial complex, patriarchal hierarchies, heterosexual orthodoxies and conventional wisdom," (Marshal 2000, p. 92), to which they had become resistant. As a result, they pushed for a more responsive curriculum, giving rise to what Pinar et al. characterize as a "crisis of meaning" (1995, p. 188). Charles Silberman's (1970) widely read indictment of education *Crisis in the Classroom* helped to usher in a new stage of curriculum development—humanistic reform. Silberman argued that "schools can be genuinely concerned with gaiety and joy and individual growth and fulfillment without sacrificing concern for intellectual discipline and development" (1970, p. 208). It was in this context that the Association for Supervision and Curriculum Development entitled their 1970 yearbook *To*

Nurture Humaneness and devoted it to how to remain humane in the midst of the broad social changes that were taking place. In one chapter of the yearbook, Francis S. Chase identifies several types of knowledge needed in order to develop "in the individual those capabilities believed to be distinctively human," including "knowledge of self," "knowledge of others," and "knowledge of the evolution and functioning of institutions" (pp. 98-100). Other curriculum texts of the period (e.g. Weinsten and Fantini's (1971) *Toward Humanistic Education*)—exhibit a similar focus on personalizing the curriculum, teaching students to retain their sense of humanity despite the pressures of modern society.

But this focus on humaneness did not last for long. By the 1980s the United States' global economic dominance was diminishing, and just as critics of the 1950s blamed lack of educational rigor for the United States falling behind in the space race, 1980s critics cited lack of educational rigor for the United States' waning economic power. In 1983 the National Commission on Excellence in Education, appointed by the Reagan administration, published *A Nation at Risk*, accusing educators of "losing sight of the basic purpose of education" and squandering "the gains in student achievement made in the wake of the Sputnik challenge" (National Commission on Excellence in Education 1983, p. 1).[6] Authors of the report call for a return the core subjects that have been neglected as a result of schools attempting to provide "solutions to personal, social, and political problems" (National Commission on Excellence in Education 1983, p. 1).

While many have since challenged the evidence used (Berliner and Bingman 1997) or the conclusions drawn (e.g. Willis et al.1993, United States Department of Education 1986) in *A Nation at Risk*, no one denies its broad impact on schools and curriculum. First, it served to widen the gap between the field of curriculum and the development of "consciously directed training experiences."[7] Second, it brought ideas like educational assessment and school/teacher accountability to the center of the conversation about American schools. These topics have retained their central position through to the present, but while many decry the current focus on standardized tests as the sine qua non of educational assessment, contemporary texts on curriculum development still include a combination of progressive and traditional themes. For example, Arthur K. Ellis divides *Exemplars of Education* (2004) into chapters on "Learner-Centered Curriculum," "Society-Centered Curriculum," and "Knowledge-Centered Curriculum" echoing the historical themes of curriculum development. In the next section, I will identify an additional theme—the co-adaptive curriculum—in order to place systems-oriented instruction within the history just discussed.

2.3 The Co-Adaptive Curriculum

With the Paleolithic educational system conveying the three most important skills—fish-grabbing-with-bare-hands, woolly-horse-clubbing, and tiger-scaring-with-fire—the tribe prospered for many years with "fish or meat for food," "hides for clothing," and "security from the hairy death." Benjamin explains, "It is to be supposed that all would have gone well forever with this good educational system if conditions of life in that community had remained the same forever" (1939, p. 33). However, the tribe did so well that after generations of fish-grabbing the slow fish available for grabbing had all been eaten, leaving only the faster, more alert fish. And after generations of horse-clubbing, the small woolly horses available for clubbing had left and been replaced by shy and speedy antelopes that could smell attackers long before they were within a club's reach. The saber-toothed tigers had become extinct (due to reasons unrelated to the scaring itself) and were replaced by bears that were not as easily repelled with fire as the tigers were.

In *The Child and the Curriculum*, John Dewey (1902) describes two camps of thought regarding education—those who support a traditional core curriculum and those who support changing the curriculum to better reflect the interests of the child. After describing each of these positions for several pages, Dewey notes, "Such oppositions are rarely carried to their logical conclusion" (1902, p. 15). While Dewey is famous for his ideas regarding educational reform, he recognized that the divide between traditionalists and reformists was not absolute. Dewey continues, "Common-sense recoils at the extreme character of each of these results. They are left to the theorists, while common-sense vibrates back and forward in a maze of inconsistent compromise" (1902, p. 15). This understanding should be kept in mind when interpreting the above history of crises. A key theme runs throughout the ebb and flow between progressivism and traditionalism: the attempt to adjust education to meet the needs of an increasingly complex society.

Before tracing the history of this theme in curriculum development, let us pause for a moment to reflect on what meeting the needs of modern society means. The systems concepts of bi-directional causality and scale may be useful here for understanding the nature of this challenge. Figure 2-1 shows a two-by-two matrix, illustrating four distinct aspects of meeting society's needs. Quadrants I and II illustrate the how society influences education. On the individual scale, schools are supposed to develop in students those skills necessary for employment and for dealing with the challenges of everyday life.[8] On a social level, this function may translate into a number of educational goals, from meeting community or national occupational needs to creating a voting public able to understand the political and social concerns of the day.[9] In quadrants III and IV the arrow of influence is reversed. In this context, schools are expected to produce students who have the ability to take a proactive role in defining both their own identity within society and redefining society to meet future unforeseen needs. Scholars may argue about where the proper balance lies in meeting these goals, but most would agree that all four are necessary to some degree. Quadrants I and II serve immediately practical goals of meeting current individual and social challenges and provide social cohesion, while quadrants III and IV enable society to grow, for each generation to find its own truths within the context of the truths it has learned. I have used the term co-adaptive in describing the curriculum to emphasize this dual nature. Responding to society's needs does not simply mean molding the individual to fit society; it must also include developing individuals who can help to mold society.

Using this model, one can trace the theme of meeting society's changing needs through the history of curriculum development crises. Even in the Yale Report of 1828, the argument for retaining the classic Latin and Greek curricula was expressed as more than a historical appreciation for our cultural heritage. Classic texts were presented as the most useful means of providing students with "the *discipline* and the *furniture* of the mind" (Italics in original text, p. 28)—i.e., providing student with thinking skills such as "the art of fixing the attention, directing the train of thought, analyzing a subject proposed for investigation" (p. 28). The focus here was on the development of the individual—quadrants I and III. While the Yale Report might be considered more readily connected to quadrant III than to quadrant I, the report's authors contended that skills acquired in the classic curricula could be generalized and applied to modern challenges. This report has been criticized not for its focus on thinking skills, but for its contention that these skills could be best developed by memorizing passages of classic texts. That is, critics agreed with the central importance of improving students thinking skills, but argued that the classic curricula were almost entirely divorced from the skills required in an industrializing nation of the 1890s.

This same criticism lies at the heart of both the turn of the century crisis that resulted in the rise of progressive education and the post WWII crisis that saw the fall of progressivism. Taba characterizes the difference between these two crises, noting that while the “criticisms of the1890s flailed against formalism, hard discipline, [and] narrowness of education,” 1950s critics faulted schools “for their softness, anti-intellectualism, progressivism, egalitarianism, [and] a lack of emphasis on fundamentals and academic skills” (1962, p. 2). But she explains that both crises were “caused by the transforming effects of technology and science on society, with criticism focusing on the failure of the schools to solve the problems created by that transformation” (1962, pp. 1-2). Similarly, Cremin suggests in 1955, “As in the period between 1893 and 1918, new social and intellectual currents are calling for new educational outlooks” (1955, p. 308). Just as traditionalists of the nineteenth century had failed to adapt to contemporary needs, the progressives of the mid-twentieth century “failed to keep pace with the continuing transformation of American society (Cremin 1961, p. 351). Tanner and Tanner echo this sentiment, citing the progressives’ “inability to recognize social change” (1990, p.262) as a primary reason for their decline.

The NEA’s 1918 report (quoted earlier)[10] illustrates how this theme of adjustment to modern needs came to be emphasized on both individual and social scales. The focus on shaping both the individual and society “toward ever nobler ends” demonstrates the inclusion of quadrants III and IV in the NEA’s view of education. Dewey’s writing during this period also emphasizes the inclusion of the social scale. Indeed, an emphasis on the impact that education can have on social development, as well as on the connections between individual and social welfare, stands as a defining contribution of Dewey's progressive ideas (e.g., Dewey 1902, 1916). While many scholars emphasize the stark differences between the progressivism before WWII and the return to a more structured core curriculum post-Sputnik, a close evaluation of the dominant texts of the late 1950s shows a continuation of this same theme of individual and social adjustment to meet modern challenges.

Smith, Stanley, and Shores' (1957) synoptic text *Fundamentals of Curriculum Development* provides a good example, not only because it was highly influential, but because their description of the social challenges of the day appear no less relevant half a century later. Their discussion rests on the premise that "the progress of science and technology has been attended by far-reaching cultural changes, which have created grave social problems" (p. 25). The authors develop this idea with a detailed account of how specialization—"minute division of labor" (p. 32)—has increased both social interdependence and individual isolation—the former because nothing gets produced without a team effort and the latter because "each individual carries around in his head a specialized picture of society, representing but a fragment of the total social pattern" (p. 32). Devoting a little space early on to bemoaning the changes at hand, the authors proceed to describing the need for "a new common sense," better suited for the realities of the day, and they are clear about the role of education in meeting this need:

> It is the obligation of those who are responsible for curriculum building to provide opportunities for children, young people, and adults to engage in the common task of rebuilding ideas and attitudes so as to make them valid for the purpose of social judgment and action in a period dominated by the complex web of impersonal relations" (pp. 52-53)

The focus here is clearly in quadrant II. And perhaps more importantly, the focus is on item A of this quadrant, a significant distinction given the charges that curriculum development during this period was focused too narrowly on the United States' need for future scientists in order to compete technologically with the Soviet Union (i.e., focus on item IIB).

Such a characterization is oversimplified. Even texts focused on the nation's occupational needs viewed the challenge more broadly. The 1958 Rockefeller Report, *Education and the Future of America*—after describing social changes similar to those identified by Smith et al. (1957)—emphasizes the need for more future scientists, but the authors explain that the challenge lies not in filling shortages in one or two occupations, but something more inclusive: "It is not a shortage now of engineers, now of economists, that lies at the root of the problem. *It is the constant pressure of an ever more complex society against the total creative capacity of its people*" (Italics in original, p. 10). The upshot of this constant pressure is clear: "Among the tasks that have increased most frighteningly in complexity is the task of the ordinary citizen who wishes to discharge his civic responsibilities intelligently" (p. 11).

If the curriculum crisis of the 1950s was brought about by the difficulty in understanding the new social complexities of the time, then the 1970s crisis was a matter of retaining a sense of personal connection to others amidst those complexities. The sense of individual isolation described by Smith et al. (1957) had increased, and curriculum scholars sought to address the problem in an affective, emotional context. Introducing the 1970 yearbook for the Association for Supervision and Curriculum Development, Scobey and Graham express their interest in "educating for humaneness during the so-called 'human revolution,' the 'technological revolution,' and the 'revolution of expanding knowledge' now in progress" (1970, p. x). Like Smith et al. (1957), contributors to Scobey and Graham (1970) identify widespread changes in society and suggest ways to adjust curriculum to best aid students in meeting new demands brought about by those changes. In terms of the educational needs model in Figure 2-1, attention simply shifted from quadrants I and II to quadrant III.

But this shift was short-lived. The Department of Education's 1983 report *A Nation at Risk* quickly steered the discussion of educational objectives back to quadrants I and II. Most of the attention received by this report has focused on the enormous influence it had on assessment and accountability in education, but the authors' expression of the social needs of the time is more pertinent to the current discussion. In the context of quadrant I, the authors explain that in an age of globalization, Americans are not only competing against other Americans, but against a global pool of potential employees and businesses. As a result, "individuals in our society who do not possess the levels of skill, literacy, and training essential to this new era will be effectively disenfranchised" (1983, p. 2). As before, technology fuels this new era, "penetrating every aspect of our lives" and "transforming a host of … occupations" (1983, p. 3). And again, the civic importance of curriculum reform is highlighted: "For our country to function, citizens must be able to reach some common understandings on complex issues," attaining "the mature and informed judgment needed to secure gainful employment, and to manage their own lives, thereby serving not only their own interests but also the progress of society itself" (1983, pp. 2-3).

My contention here is not that these historical texts are all calling for the same thing. Indeed, one could hardly find a pair of texts that differed more than *To Nurture Humaneness* and *A Nation at Risk*. The disparity between these texts only strengthens the point at hand: Historical calls for educational reform, for all of their differences, have shared a focus on adjusting curriculum to meet the increasingly complex demands of an increasingly complex society. In the following section, we will look at our contemporary social needs and how they translate to educational goals.

2.4 A New Crisis?

As a result of the prehistoric changes, the Paleolithic tribe no longer had food, clothes, or security from the hairy death. But in a short time, the Paleolithic tribe's innovators caught up with the changes with the fish, the antelope, and the bears. Fishers learned to net fish rather than grab them, horse-clubbers learned to snare antelope instead of clubbing them, and tiger-scarers learned to dig bear pits instead of using fire. As a result, the tribe had more fish, meat, and skins than they had ever had before. Some suggested that in light of these new conditions the educational curriculum be adjusted to address these new skills. But the majority argued that the curriculum was already filled with fish-grabbing, horse-clubbing, and tiger-scaring, leaving no room for "fads and frills like net-making, antelope-snaring, and—of all things—bear killing" (Benjamin 1939, p. 43).

In light of the history recounted above, one might be hesitant to declare yet another crisis in education. Given the luxury of historical perspective, one can view the series of crises as simply the continuing development of curriculum necessary to meet the changing demands of a changing society. In this context, it may be worth surveying some contemporary changes in society. While the empirical research in this report has been designed in the context of social challenges and civic literacy—quadrants II and IV—much of the promise of systems-oriented instruction lies in its applicability to quadrant I. The hypothesis in the empirical portion of this report is that systems-oriented instruction better prepares students for understanding contemporary social and ecological challenges. However, systems-oriented instruction is not simply environmental education. Before discussing larger social challenges, it is worth taking a brief look at the employment demands on today's student.

Describing the collapse of progressive education after WWII, Cremin explains, "The economy had entered upon an era marked by the harnessing of vast new sources of energy and the rapid extension of automatic controls in production, a prodigious advance that quickly outmoded earlier notions of vocational education" (1961, p. 351). A recent report from the United States Department of Education echoes this same shift: "Whether filling 'blue collar' or 'white collar' positions, employers seek…practical problem-solvers fluent in today's technology. If current trends continue, by 2012, over 40 percent of factory jobs will require postsecondary education" (2006, p. 4). In this context, being a practical problem-solver means more than being technologically savvy. In a 1989 article in *Fortune Magazine* Brian Dumaine, describes a new trend in American industry, explaining that "the most successful corporation of the 1990s will be something called a learning organization, a consummately adaptive enterprise with workers freed to think for themselves, to identify problems and opportunities, and to go after them" (p. 48). A joint report produced a decade later by the United States' Departments of Commerce, Education, and Labor supports Dumaine's assertion, predicting that economic success in the 21st century "will require adopting organizational work systems that allow workers to operate with greater autonomy and accountability" (1999, p. iii). Table 2-1, taken from the 1999 report, illustrates the organizational shift from linear hierarchies to flexible networks that will require employees to have a broader understanding of their organization's operations (U.S. Department of Commerce and U.S. Department of Education 1999, p. 3).

One can best understand the significance of this table in the context of the distinction Peter Senge (1990) makes between detailed and dynamic complexity. Detailed complexity refers to a system with many variables. The task in a large high school of meeting as many as 1600 students' scheduling needs and preferences within the constraints set by class size, classroom space, and teacher schedules involves a high level of detailed complexity, but low dynamic complexity. That is, the problem involves far too many variables for an administrator to manage in his head, but the variables do not interact with one another. Dynamic complexity, conversely, may involve relatively few variables. The complexity arises as a result of the interactions between the variables.

The most common example of this in an ecological context is the Lotka-Volterra predator-prey model, which includes only one prey species and one predator species. Even in this relatively simple model, involving only linear equations to define the relationships between the populations of the two species, the interactions between the species result in dynamic oscillations of their populations over time (See Figure 2-2). Senge (1990) uses the example of the cold-war arms race, which, even simplified to only six variables, possesses high dynamic complexity due to the interactions between the variables. Senge's (1990) example also illustrates how dynamic complexity, far from being a purely academic concept, can play a crucial role in our everyday lives. We will return to the example in the following chapter. For now, the main point is simply that in today's workplace, dynamic complexity has come to play a larger role than in the past.

But today's students will face a number of new challenges as well. Recall that in the history of educational crises, calls for curriculum change were generally made in the context of social change. That is, observations regarding increased complexity referred primarily to changes in relationships between people, often brought about by technological advance. Today, in addition to changing the nature of our relationships with each other, technological advance has also changed our relationship with the environmental systems that support us. Ervin Laszlo alludes to this change in his discussion of the logic of the modern industrialized society—"a logic," he explains, "that led from the 'progressive appropriation of the world' to 'progress that masters the world,' all the way to the environmental, economic, social and cultural limits inherent in [industrialization]" (Laszlo et al. 1993, p. xvi). In other words the complexity inherent in earlier stages of industrialization involved learning to succeed in a world made faster, busier, less personal, and more complicated by increased appropriation of resources, but the resources themselves were infinite, as was the ability of the earth to deal with the wastes produced through the use of those resources. The complexity cited in the early quotations above still refers to a world without limits. Today, the idea of a "complex society" has grown to include the limits of a finite world, and as we approach these limits, Laszlo suggests, "the developmental curve of modern industrial society registers a turnaround" (Laszlo et al. 1993, p. xvi).

One need not look very hard to observe evidence of this turnaround. For example, the earth currently holds 6.3 billion people, roughly twice the population of the 1960s when human effects on environmental systems first became widely noticed in the United States, and that number is expected to grow to almost 9 billion by 2050 (Cohen 2003). In addition, the level of consumption per capita in industrialized nations has increased. In the United States personal consumption expenditures increased 33% from 1993 to 2004 (Council of Economic Advisors 2005).

Increased human population and consumption has led to increased problems managing wastes from human activity. For example, industrial air pollution now poses a serious health risk on both regional and global scales (Akimoto 2003; Ezzati et al. 2005), and global climate change has now been officially recognized by leading nations, including the United States, as a "serious and long-term challenge" attributed "in large part" to human activities (G8 Gleneagles 2005). Add to these concerns the loss of biodiversity (Jenkins 2003), the collapse of large fisheries (Pauly et al. 2003; Essington et al. 2006), the loss of forests and soils (Stocking 2003; Wright and Muller-Landau 2006), and the projected scarcity of potable water (Tully 2000; Gleik 2003), and one might be tempted to look nostalgically back at, say, the 1950s, when an increase in the complexity of society referred *only* to the dehumanization of the individual as a result of industrialization.

Ecologists Howard Odum and Elizabeth Odum (2001) write extensively about this turnaround in the text *A Prosperous Way Down*. The Odums argue that, as their title suggests, the prospect of making the adjustments necessary to address these challenges need not inspire apocalyptic visions of collapse. In their view "the global society can turn down and descend prosperously, reducing assets, population, and unessential baggage while staying in balance with its environmental life-support system" (2001, p. 3). However, such a path would require significant changes in social policies and institutions analogous to those outlined in Table 2-1.

Therefore, without belittling the problems of prior generations, it seems fair to say that today's students will face a set of challenges qualitatively different than those facing students when Smith et al. wrote:

> [An educational program must include] new patterns of thinking, wherein a number of social variables in politics, economics, and the like are kept in the picture in the process of reaching conclusions about social policies and actions, instead of the prevailing and now obsolete habit of thinking in a linear and compartmentalized fashion (where, for example, the attempt is made to keep political and economic thought in their separate spheres). (1957, pp. 95)

We can nevertheless learn from their wisdom, continuing to ask ourselves what new patterns of thinking might best enable us to address contemporary challenges.

2.5 Conclusion

The story of the Paleolithic tribe struggling with its own curriculum controversies (e.g., tiger-scaring versus bear-pit-digging) ends with the Paleolithic youths, bored by the obsolete curriculum, becoming listless underachievers. Meanwhile, a neighboring tribe, a thinly veiled version of Hitler's Germany that has taken a more pragmatic position regarding curriculum affairs, invades. The story, written in 1939, is in some ways a product of its time, but one need not try very hard to see echoes of it in the stories implied by George W. Bush and David W. Orr at the opening of this chapter. In Bush's version, the neighboring tribe would not be Nazi Germany, but perhaps China or India. And the invasion would not be military, but economic. In Orr's version, the invading tribe would not be people at all, but rather the compounding problems created by our own environmental negligence collapsing upon us. In any case, the point would be the same: a curriculum must evolve with the society for which it is designed.

In this chapter I have reviewed the history of curriculum development, characterizing it as a series of curriculum changes brought about through co-adaptation with changes in society itself. The practice of adjusting curriculum in order to address the educational needs of an increasingly complex society is nothing new. However, the degree of complexity has grown exponentially with our population and our technological capability to affect large-scale systems (e.g. climate change).

In 1861 Herbert Spencer asked a question that has been quoted so often in curriculum texts, one might consider it the north star of the field of curriculum development: "What knowledge is of most worth?" (p. 11). We might also take Smith et al.'s lead and ask, What patterns of thought are of most worth? Many scholars, scientists, educators, business leaders, and politicians have suggested that systems thinking represents a set of skills particularly well suited for addressing our current challenges. In the following chapter, I will describe what is meant by systems thinking and discuss how it might be applicable to our current situation.

Notes

[1] See also Cremin (1971).

[2] Kliebard (1995) reports of a 1913 survey of child laborers in Chicago. Helen M. Tood, a factory inspector asked 500 children working in the factories if they would prefer to be in school if their families could afford it. Kliebard

explains, "Of the 500, 412 told her, sometimes in graphic terms, that they preferred the often-grueling factory labor to the monotony, humiliation, and even sheer cruelty that they experienced in school" (6).

[3] In *Democracy and Social Ethics*, Jane Addams provides many examples of how the industrial work could be humanized by providing future workers with a broader perspective of the system. For example, she explains, "It takes thirty-nine people to make a coat in a modern tailoring establishment, yet those same thirty-nine people might produce a coat in a spirit of 'team work' which would make the entire process as much more exhilarating than the work of the old solitary tailor" (1902, 219).

[4] The current Bush administration cites the success of this program as a model for the American Competitiveness Initiative (U.S. D.O.E 2006, p. 4).

[5] In addition to being associated with a back-to-the-basics movement, Bruner is also criticized for removing the curriculum specialist and teacher from the process of curriculum development. The almost total absence of teachers and curriculum specialists at the Woods Hole conference lends credence to this criticism.

[6] George Willis et al. (1993) note a key difference between this and the post-Sputnik efforts of the 1960s: "In that era the federal government began to supply large amounts of funding for schools; in the 1980s, in contrast, the federal government placed responsibility for funding educational reforms on state and local governments while at the same time reducing the overall level of revenues it dispersed to states and local communities" (401).

[7] Recall this term from Bobbit's duality explained in the opening of this chapter. In the 1970s and 1980s the field of curriculum moved away from planning the classroom experience and focused instead on understanding aspects of the "hidden" or unofficial curriculum. This scholarship involves deconstruction of the curriculum in order to identify hidden power relationships with an eye toward, for example, feminist, racial, and literary theory (e.g. Pinar 1998, McCarthy and Crichlow 2005).

[8] Philosopher Richard Rorty explains, "Even ardent radicals, for all their talk of 'education for freedom', secretly hope that the elementary schools will teach the kids to wait their turn in line, not to shoot up in the johns, to obey the cop on the corner, and to spell, punctuate, multiply, and divide" (1999, 117).

[9] Large scale social norms, such as valuing the tenets of a democratic government, might also fit into this quadrant, but would overlap into quadrant IV.

[10] "[E]ducation in a democracy…should develop in each individual the knowledge, interests, ideals, habits, and powers whereby he will find his place and use that place to shape both himself and society toward ever nobler ends." (Department of the Interior 1908, 157)

Table 2-1. Jobs are changing due to shifts in organization and management

Element	Old System	New System
Workplace organization	Hierarchical Rigid Function/specialized	Flat Flexible Networks of multi/cross-functional Teams
Job Design	Narrow Do one job Repetitive/simplified Standardized	Broad Do many jobs Multiple responsibilities
Employee Skills	Specialized	Multi/Cross-skilled
Workforce Management	Command/control systems	Self-management
Communications	Top down Need to know	Widely diffused Big picture
Decision-making responsibility	Chain of command	Decentralized
Direction	Standard/fixed operating Procedures	Procedures under constant Change
Worker autonomy	Low	High
Employee knowledge of organization	Narrow	Broad

	Individual	**Social**
Society→Education	**I.** A. Personal skills for everyday challenges B. Skills for employability	**II.** A. Ability to understand current social needs B. Ability to meet national occupational needs
Education→ Society	**III.** A. Ability to forge one's sense of individuality B. Inclination toward personal development	**IV.** A. Ability to perceive and respond to future social needs B. Ability to develop and apply innovative means of improving society

Figure 2-1. Four aspects of meeting the educational needs of society

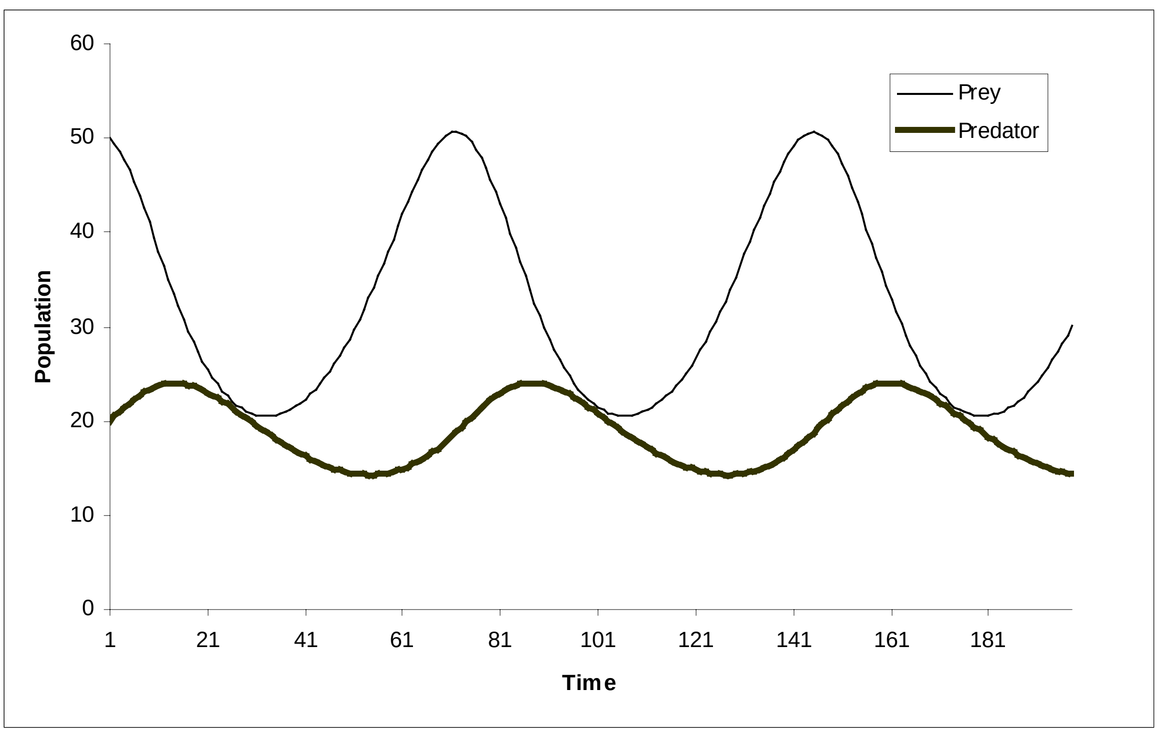


Figure 2-2. Oscillations in the Lotka-Volterra predator-prey model

CHAPTER 3
SYSTEMS-ORIENTED INSTRUCTION

> A systems viewpoint is not necessarily a better one than any other, just a different one. Like any viewpoint, like the top of any hill you climb, it lets you see some things you would never have noticed from any other place, and it blocks the view of other things. Systems training has taught us to see the world as a set of unfolding dynamic behavior patterns, such as growth, decline, oscillation, overshoot. It has taught us to focus on interconnections. (Meadows *et al*. 1972, p. 2)

3.1 Beyond Reductionism

In the previous chapter I argued that conventional curricula fail to adequately prepare students for the challenges they face in an industrialized, globalized world. I discussed these challenges in the context of both employability and civic literacy, and suggested that systems-oriented instruction may represent a significant improvement in preparing students for the complexity of the 21st century. In this chapter I will describe what I mean by systems-oriented instruction, including both a theoretical discussion and practical examples of specific tools used in the classroom. I will conclude this chapter with a discussion of why systems-oriented instruction shows promise in the context of the educational challenges discussed in Chapter 1.

There is nothing intrinsically wrong with the structure of conventional curricula commonly used at primary, secondary, and tertiary levels of education. Conventional curricula—indeed, the structure of the schools themselves—have been developed in a reductionist context and, consequently, come with all the advantages and disadvantages of reductionism. Often associated with the ideas of Isaac Newton and Rene Descartes, reductionism can be defined as the belief that one can best understand an object of study by understanding its component parts. Thus, in order to learn about something, one must take it apart and learn about its parts. In order to learn about those parts, one must take them apart and learn about the parts' parts. And so it goes until one has reached some fundamental set of components that can no longer be dissected. While discoveries in quantum physics have muddied the idea of fundamental components, one can still see the common sense of such a view. In order to learn about a car engine, it makes sense to look under the hood and start tinkering, and one can learn much about the anatomy of a frog by dissecting it. But this type of study would tell you little about the effect that ownership of a car might have on the lifestyle of an individual or about the frog's role in an ecosystem.

In short, while analyzing an object of study part-by-part can tell one much, such analysis will likely fail to reveal important aspects of the object. This insight represents the fundamental shift for proponents of the position identified as systems science or general systems theory. Scholars often refer to the shift to a systems worldview as a paradigm shift, alluding to Thomas Kuhn's *Theory of Scientific Revolutions* (1987), but such reference can be misleading, as the term "paradigm shift" often refers to the adoption of a new system of beliefs that are irreconcilable with the prior system. The example most commonly cited is the shift from Ptolemaic to Copernican astronomy. Kuhn explains that before this shift "the sun and moon were planets, the earth was not. After it, the earth was a body, a satellite. Changes of that sort were not simply corrections of individual mistakes embedded in the Ptolemaic system" (Kuhn 1987, p. 8). In other words, adopting the Copernican system meant the wholesale rejection of the Ptolemaic one. The shift to a systems view is not quite so all-encompassing.

Proponents of systems theory do not deny the utility of reductive analysis; they simply deny its primacy. Even the most ardent critics of reductionism recognize the amazing scientific advancements made during what Ackoff (1999) refers to as the Machine Age. These advancements are far too numerous to mention here, but articles in the January 2006 issue of *Discovery Magazine*—including such subjects as carbon nanotubes, which are nine times stronger than steel and one thousand times more conductive than copper; laser transistors that promise to operate at about 100 times faster than the top-of-the-line personal computer chips available today; and a vaccine for cervical cancer—attest to the success of reductive science. And yet, reductive science has often been found to be lacking, particularly in the context of living systems. Ackoff's label carries a double meaning. Not only has the machine age seen incredible advancement with regard to building more and more complicated machines, but it has come to use the machine as the primary metaphor for understanding the world.

From a systems view, the machine metaphor often leads one to overlook aspects of the object of study that systems proponents consider to be significant, particularly in the context of living systems. Biologist Paul Weiss explains, "We can assert definitely…on the basis of strictly empirical investigations, that the sheer reversal of our prior analytic dissection of the universe by putting it back together again, whether in reality or just in our minds, can yield no complete explanation of the behavior of even the most elementary living system" (Weiss 1971, p. 267; Quoted in Capra 1982, p. 102). For systems thinkers like Weiss, learning how the parts behave individually does not tell one how they will behave together as a whole. One must study the pattern of interactions between the parts to gain a better understanding. In this context, Gregory Bateson (1987, p. 153) characterizes a systems view as a shift from the nouns (things) to the verbs (interactions).

This seemingly subtle shift of focus had important consequences in a number of different fields, including psychology, sociology, biology, and cybernetics, just to name a few. Therefore, the shift from a reductive to a systems view can be considered a paradigm shift in the sense that it involves fundamental changes in how one approaches problem solving. As Ackoff explains, "Systems is more than just a concept. It is an intellectual way of life, a worldview, a concept of the nature of reality and how to investigate it" (1999, p. 1). The following section provides more detail regarding how one might investigate reality from a systems perspective.

3.2 A Systems Perspective

Ludwig von Bertalanffy, a founding father of systems theory, defines a system mathematically as a set of elements, where "elements, p, stand in relations, R, so that the behavior of an element p in R is different from its behavior in another relation R' " (1969, pp. 55-56). For now, von Bertalanffy's less formal definition should suffice: "a group of elements standing in interrelation with themselves and with the environment" (1969, p. 252). The language is intentionally general. Examples of systems can range anywhere from a single cell to the earth as a whole. Note the key role of relationships in both of von Bertalanffy's definitions. The central idea of systems theory is that relationships between elements result in the emergence of properties and behavior that do not exist in isolated elements.

The significance of this insight depends on the system one is studying. Ackoff (1999) distinguishes between four kinds of systems—deterministic, animated, social, and ecological—based on the role that purpose plays in each system.[1] In deterministic systems, "neither the parts nor the whole are purposeful" (e.g. a computer); in animated systems, "the whole is purposeful but the parts are not" (e.g. a person); in social systems, "both the parts and the whole are purposeful" (e.g. a community); and finally, in ecological systems, some of the parts are purposeful, but not the whole (e.g. a wetland) (Ackoff 1999, p. 27). One can think of these types in the context of the distinction from Chapter 2 between detailed and dynamic complexity. Deterministic systems may have a high degree of detailed complexity, but will have a low degree of dynamic complexity. For example, a personal computer or a car each have a large number of interacting parts, but those parts are highly limited by their structure. It is worth noting here that interactions between parts are not entirely ignored in the reductionist framework. Billiard balls colliding with each other are indeed interacting. However, the interactions are severely limited and entirely deterministic. The reductionist paradigm, for the most part, provides an adequate context for studying deterministic systems. Consequently, most of the technological advancements of the Machine Age have been in the context of deterministic systems.

Conversely, animated systems, which for Ackoff contains most animals including humans, can, and often do, have a high degree of dynamic complexity. In these systems interactions between parts and between parts and the whole are far more variable. While physical laws still apply, the responses of these systems are not defined solely by those laws, but also by an element of choice.[2] Gregory Bateson (1979) dryly illustrates this distinction by describing the difference between kicking a stone and kicking a dog. The response of the stone is far more predictable[3].

Ackoff's typology can become problematic if interpreted as strict divisions, since different species exhibit behavior with incrementally different levels of determinacy. However, interpreted as a spectrum, Ackoff's typology becomes a useful tool for understanding the limits of reductionism. The further a system moves beyond deterministic behavior, the less suitable a reductionist model will be for describing that system. Since Ackoff's last two categories—social systems and ecological systems—both involve multiple animated systems interacting with each other, the level of dynamic complexity is generally quite high. That is, the patterns of interaction take an even larger role in the dynamics of the system as a whole. Building upon Bateson's example, one can imagine how a pack of dogs might exhibit behavior that none of the dogs would exhibit individually. In the following four sections we will look at some basic concepts useful for learning about dynamically complex systems.

3.3 Nonlinear Causality

At the heart of systems thinking lies an understanding of nonlinear causality. That is, shifting from the fundamental model of a causal chain to that of the causal web. Figure 3-1 illustrates this difference. The first insight gleaned from this shift is that there is not necessarily a one-to-one correspondence between causes and effects. Using Figure 3-1a as a model, one might suggest that a change in element A would lead predictably down the causal chain to effect a change in element E. However, thinking of the elements in terms of the causal web in Figure 3-1b makes one more aware of other causal relationships that may have been overlooked in the causal chain. First, notice that element B is influenced by both elements A and L. Thus, the role that element A plays on B may not be as clear as indicated in the causal chain. Secondly, notice that element A also influences element F, which in turn influences G and H. Thinking of events in terms of causal chains can lead to tunnel vision—a focus on one set of events to the point that other causal connections are ignored. We can express this first insight in the following two rules: (1) One cause can have more than one effect, and (2) one effect can be the result of more than one cause.

The second insight has to do with what we might call dynamic interaction. In dynamic interactions, the interaction between any two elements is dependent upon other interactions taking place at the same time. To understand the significance of this insight, let us return briefly to the billiard ball world of reductionist science, where all cause and effect relationships can be isolated so that one can study the relationship between one cause and one effect. One can calculate the effect of a cue ball hitting the eight ball, given the velocity and mass of each ball. If two balls are hitting the eight ball at the same time, one can mathematically isolate the two collisions and calculate the exact role that each ball played in the subsequent movement of the eight ball. This method of isolation and analysis will work for any number of billiard balls when one assumes perfect elasticity and neglects friction and gravitational pull between the balls. That is, increasing the number of balls involved increases the detail complexity of the system, but not the dynamic complexity.

By neglecting gravitational pull between the billiard balls, one is essentially assuming that the balls are not interacting except at the collisions. That is, one is ignoring the interaction between the balls at any other time. For billiard balls, the assumption makes sense since their masses are so small. In systems with larger bodies, planets for example, gravitational pull between the bodies plays a significant role. The necessary inclusion of the gravitational interaction between the bodies changes the nature of the problem, making it impossible to mathematically define the motion of even three bodies interacting with each other.[4] This difficulty arises because the gravitational force both effects and is dependent upon the position of the bodies. In the context of our systems language, we can say that the inclusion of gravity into the problem increases the dynamic complexity of the system, making ordinary analysis less applicable. Here, of course, we are talking about a purely physical system. The nature of interaction between bodies (or elements) becomes more variable in the context of living (animated) systems.

Returning to Figure 3-1b, we can now see that element B may depend on a combination of elements A and L. That is, perhaps neither element A nor element L would bring about element B individually, but the dynamic interaction of both elements A and L will result in B. The term *emergent property* is used to define those properties in a system that cannot be observed in any one element in the system but emerges as a result of interaction between elements.

Perhaps the most common illustration of these first two insights comes from medication advertisements on television. These commercials generally follow the same structure. In a thirty-second spot, the first twenty-five seconds are spent describing or illustrating the desirable effects offered by the medication. The last five seconds are spent expressing two types of warnings. First, the viewer is bombarded with a rapid-fire list of side effects associated with the advertised medication. In other words, one event—taking the medication—leads to a number of effects. In this context, the term *side effect* is purely subjective. The desired outcome from the medication and the dry-mouth, headache, etc. that also occur are all simply effects of the medication. We use the term *side effect* to identify those effects that are unintended and often not particularly desirable. Second, the viewer is warned not to mix this medication with certain other kinds of medications because the combination may produce effects that neither medication would have produced on its own—two events combining to produce an effect that neither would have produced alone.

Returning again to Figure 3-1b, we can see another type of dynamic interaction. Note that in addition to element A influencing element B, element B influences element A. Similar to the planetary example, the influence that element A has on B will depend on the influence that element B has on A. We can see this same phenomena occurring indirectly with elements A, B, C, and I and with elements C, D, and E. Each of these sets of element forms what is called as *causal loop*, where the influence of an element indirectly affects that element itself. Element C affects D, which affects E, which comes back to affect C. One cannot overstate the importance of the role causal (or feedback) loops play in the often counterintuitive behavior of complex systems. Therefore, discussion of these loops comprises the following two sections.

3.4 Causal Loops

In a recent interview about global warming, Al Gore explained the dynamics of the polar ice cap melting in the Arctic Ocean: “When the ice there melts, there’s a dramatic change in the relationship of the surface of the earth there to the sun. The ice reflects 90% of the incoming sun’s energy like a mirror, but the open seawater, after it melts, absorbs 90%. And that’s a phase change. It sets up a *positive feedback loop* that magnifies and speeds up the melting process” (Gore 2006, italics added). Here we have an example of a well-known politician employing a systems concept on mass media in order to explain the details of global warming. The problem is that for most people the term *positive feedback* refers to someone saying “Good job” after a piano recital or some other performance.

Gore means something quite different. To understand the systems notion of positive feedback, we must first understand the notion of a causal loop. Since most people have a tendency to view events as causal chains, Figure 3-2a expresses most of what Gore describes as a causal chain. This chain of events occurs, as Gore explains, because liquid water absorbs more of the sun’s energy than ice. But to understand the full impact of Gore’s point, one needs to view this chain of events as a loop of events as in Figure 3-2b. Looking at the events as a loop, one can understand the ongoing effect of the melting ice cap. As the ice melts, the Arctic Ocean absorbs even more energy, causing more ice to melt, causing the absorption of more energy, and so on. Thus, a small change in the polar temperatures can be amplified by this cycle over time. This type of feedback loop is called a positive (or reinforcing) feedback loop because of how it amplifies small changes.

Peter Senge (1990) illustrates how positive feedback loops can play a role in politics as well, using one of the most widely known examples of a reinforcing loop: the arms race. Senge explains, “The roots of the arms race lie not in rival political ideologies, nor in nuclear arms, but in a way of thinking both sides have shared” (1990, p. 71). Figure 3-3a and 3-3b show how Senge represents both the American and the Soviet perspectives of the arms race. Senge notes that both Americans and Soviets view each other as the aggressor and themselves as simply responding to a threat. Figure 3-2c illustrates the systems view of this dynamic, combining the two chains into one loop with reinforcing or positive feedback. From this perspective, one can see the cold war as a quintessential example of a positive feedback loop. Note that unlike the more casual use of the term “positive feedback,” the use of the term as a systems concept has nothing to do with the desirability of the change being amplified. In fact, the majority of positive feedback loops identified in the literature describe undesirable changes, just like the two examples here. The term “viscious cycle” has been coined to describe these reinforcing loops that amplify undesirable changes.[5]

The second type of feedback loop is called a negative (or balancing) loop. A household thermostat is the example most commonly used to illustrate this type of feedback loop. When one sets the temperature on a thermostat, one is not setting a point temperature but rather a range of acceptable temperatures. In the summer, one’s air conditioning will turn on as soon as the house temperature has risen above the upper boundary of the set range. The air conditioner will continue to cool the house until the house temperature drops below the lower boundary of the set range, at which point the air conditioner turns off. Figure 3-4 illustrates the repeating cycle of events. Notice that unlike a positive feedback loop, events in this loop work to counter change, not reinforce it. Other examples of balancing feedback loops include our own biological thermostats by which we maintain a relatively constant body temperature and the predator-prey system described in Chapter 2, where the populations of both the predator and the prey oscillated within a specific range.

A complex system can have any number of positive and negative feedback loops, and almost any combination is possible. However, there are a handful of typical combinations of these loops that seem to appear quite often in a number of different contexts. An understanding of these most common configurations can help one to understand the often puzzling behavior of complex systems. Section 3.6 focuses on these configurations, but since these configurations are generally expressed as diagrams, a brief lesson on diagramming is in order. The following section provides this lesson.

3.5 Identifying Positive and Negative Feedback in Systems Diagrams

Systems can be diagramed to illustrate quantitative relationships, expressing specific values associated with each stock and flow. However, they can also be simplified to show qualitative relationships, where the arrows are labeled only with + and -, instead of with specific numbers, to illustrate the nature of the relationship between the two nodes. Labeling an arrow with a + implies that the two nodes move in the same direction. An increase in the causal node results in an increase in the effect node, and a decrease in the causal node results in a decrease in the effect node. This describes all of the relationships in Figure 3-3c, so we can redraw the arms race diagram more formally by labeling the arrows as in Figure 3-5.

A negative sign is used to illustrate when the two nodes in question move in opposite directions—an increase in the causal node results in a decrease in the effect and a decrease in the causal node results in an increase in the effect. Such a relationship exists between the predators and prey in our simplified ecological example. An increase in the predator population leads to a decrease in the prey population as the predators consume more and more prey. However, a decrease in the prey population implies a lack of food for the predators, leading to a decrease in the predator population. These relationships may sound complicated, but they can be expressed in a diagram with just two stocks and two flows, as in Figure 3-6.

Once one has accurately labeled each flow with + or -, identification of positive and negative feedback loops becomes a matter of simple integer arithmetic. The polarity of a loop can be calculated just as if one was multiplying integers. The loop in Figure 3-5 contains only positive arrows. Multiplying a positive integer by a positive integer yields a positive integer. Therefore, the loop is a positive feedback loop. A small change in any of the variables will be magnified as the change runs through the causal loop. Conversely, the loop in Figure 3-6 contains one positive arrow and one negative arrow. Multiplying a positive integer by a negative one yields a negative integer. Therefore, the loop in Figure 3-6 is a negative feedback loop. With this understanding we are now ready for a discussion of the common systems configurations mentioned in the previous section.

3.6 Systems Archetypes

Early in the 20th century, Carl Jung introduced the term *archetype* to psychology to signify those corresponding themes or ideas that he identified in dreams, myths, fairy tales, and tribal lore.[6] For Jung these themes transcended culture, providing a fundamental psychological template that shaped human perception of experience. In a literary sense, we might think of archetypes as generalized descriptions of characters that recur in mythology and literature. Proponents of systems thinking use a similar term, *systems archetypes*, alluding to Jung's sense of universality. We can think of systems archetypes similarly as basic structures or configurations recurring in complex systems in a wide array of differing contexts.

The specific number of archetypes considered to comprise the basic systems toolkit ranges from eight to twelve. My purpose here is not to provide a comprehensive discussion of systems archetypes, but rather to introduce the concept and illustrate its utility as a pedagogical tool. Therefore, I will only discuss a few of the most common archetypes. We are already familiar with the two most basic archetypes: reinforcing loops and balancing loops. The generic diagrams for these archetypes are shown in Figure 3-7.

A third archetype, called Escalation, is a specific kind of reinforcing loop, which again, we have already seen. In Figure 3-5 I show the arms race example as one large loop in order to make the reinforcing cycle clear. However, this diagram can also be drawn to emphasize how each party is focused on its own balancing loop, as in Figure 3-8. The generic Escalation diagram is shown in Figure 3-8a. Notice how the two loops combine in figure-eight fashion to create a positive feedback loop like that in Figure 3-5. Figure 3-8b shows how the Cold War example can be diagrammed in the Escalation form.

A fourth archetype, Shifting the Burden, depicts our tendency to focus on symptoms of problems rather than addressing the problems themselves. As Figure 3-9 shows, this archetype contains two balancing feedback loops. In one of these loops, the actions taken to address the problem seem to be having a beneficial affect. In truth, however, the actions fail to address the underlying cause of the problem, so the problem returns, often worse than before. Figure 3-9b illustrates the dynamic documented by Jason Evans (2006) regarding the water hyacinth in Florida springs. Water hyacinth is an attractive, but exotic species that thrives in nutrient-rich waters. As development within the watershed of a spring increases, the nutrient level coming up through the spring also increases, resulting in an explosion in the water hyacinth population. The policy for addressing this problem has often been to remove the hyacinth from the springs and deposit the plants onshore to die. While this solution does temporarily decrease the water hyacinth, it does not address the high nutrient level in the water that caused the hyacinth to grow so quickly. The result is often repeated hyacinth blooms or a bloom of lyngbia, a toxic but native algae. Evans (2006), addressing the problem from a systems perspective, has focused on the nutrient load itself and suggested using the water hyacinth as a tool for managing nutrient levels in the water.

The last two archetypes we will discuss here—Limits to Growth and Tragedy of the Commons—both received their names from classic natural resource texts. The Limits to Growth archetype outlines the dynamics described in the Meadows *et al.* (1972) text of the same name. The main idea for this archetype, diagrammed in Figure 3-10, is that while a variable may for a short time grow exponentially through a reinforcing feedback loop (on the left of the diagram), its growth will be limited by a larger-scale balancing feedback loop (on the right of the diagram). At this point the variable that had been growing will decrease to a point within the limits of the balancing feedback loop. In a business context, this archetype is used to help managers learn to look ahead for factors that will potentially limit the growth of the business and to invest to address those limits. William Braun (2002) cites the example of the Internet provider America OnLine. Initially, America Online invested heavily in an aggressive marketing campaign, inundating potential subscribers with their software. The campaign was extremely successful, and America Online experienced rapid growth in subscribers. However, growth occurred so quickly that their system was overwhelmed, resulting in slower service, which in turn led to decreased subscriptions. In this case, system capacity was the limiting factor.

In their text, *Limits to Growth*, Meadows *et al.* (1972) focus not on removing limits to growth, but rather on working within those limits that cannot be removed. Specifically, they investigate how growth trends in population, agricultural production, and industrialization will be limited both by depletion of natural resources and by the limits of the environmental systems on which we depend to process the waste materials that accompany current agricultural and industrial processes. The report evoked strong responses, largely because people believed—or wanted to believe—in unlimited growth of current trends and practices. The details of the report have since been amended (Meadows *et al.* 1972), but its exploration into how natural resource limits will affect global development remains a valuable lesson.

In his article "Tragedy of the Commons," Garrett Hardin (1968) introduces another valuable resource lesson, illustrating how rational behavior on the individual scale can lead to a collective condition that hurts everyone involved. The diagram for the archetype, Tragedy of the

Commons (Figure 3-11a), looks complicated at first glance, but focus first on the reinforcing feedback loop in the upper left of the diagram. Participant A receives benefits from an activity. These benefits allow A to increase its activity and gain even more benefits. Participant B experiences the same thing in the reinforcing loop on the lower left. Thus, each individual, focused on his own reinforcing loop, continues to increase his activity. But problems occur when the combined actions of the participants degrade the resource, eventually leaving it unable to support either participant's activities. Figure 3-11b shows how this archetype looks in the context of two participants whose combined efforts overwhelm a fishery, but it can involve any number of participants and can describe any system involving a common resource that can be overwhelmed by overuse. A large company's IT or human resource department can be overwhelmed by other departments each focused on their own tasks rather than the collective workload; electricity demands on a summer afternoon resulting from wide use of air conditioning can cause a blackout; and individual family members' desire for a hot morning shower can leave everyone standing under cold water. All of these situations can be described by the Tragedy of the Commons archetype.

Examples of each of these archetypes—and those not discussed—abound in social, economic, and ecological systems. Thus systems archetypes provide examples of how systems-oriented instruction could become a powerful pedagogical tool for promoting the transfer of understanding from one system to another. For example, a common systems lesson for primary students involves a systems explication of Dr. Seuss's *The Butter Battle Book*. You may recall that this text documents the dispute between the Yooks and the Zooks over the proper side on which to butter one's bread. The dispute leads to an arms race that begins with a slingshot and ends with each side threatening to use a Big Boy Boomeroo, a sort of ultimate weapon capable of destroying both societies. Even young students can understand the Escalation dynamic at play here when led through a diagramming exercise that results in a diagram analogous to Figure 3-5.

It is perhaps not hard to imagine how, after working through such an exercise, students might have an easier time understanding the basic dynamics behind how a schoolyard shoving match can lead to a shooting, why countries trying to attract businesses adopt looser and looser environmental and labor regulations, and perhaps even why Gillette recently unveiled a five-bladed razor—the Fusion—reportedly besting the competition's four-bladed models. Similarly, a student familiar with Shifting the Burden might be better able to understand why building more roads results in more traffic, why drug use is not a successful problem-solving technique, and why teaching a man to fish is better than giving him one. The list of examples could go on and on.

Of course, real systems are often more complex than the generic archetypes described above. True understanding of a system would require more than simply forcing it into a rigidly predefined shape. The point is that the systems concepts and archetypes discussed in this chapter can provide valuable tools both for transferring learning from basic to more complex systems as a student matures, but also from systems in one field of study to another. The following section addresses the importance of both types of transfer.

3.7 Systems Thinking and Transfer of Learning

The distinction between systems-oriented instruction and conventional instruction can best be understood in the context of what Chris Agyris calls "double-loop learning." In single-loop learning, individuals learn new facts, which they interpret and incorporate into their mental models. Double-loop learning involves restructuring those mental models, thereby altering the way individuals interpret new information. Shifting to systems-oriented instruction does not mean teaching different facts; it involves teaching different mental models for interpreting facts. It is in this context that systems thinking is often referred to as a critical thinking skill.

A systems-oriented curriculum can include all the same subjects employed in conventional curricula (math, social studies, language arts, etc.). However, in systems-oriented curricula, each of these topics is approached in the context of systems concepts, resulting in more specific attention to the connections between the various subjects. Thus the same tools and concepts used to understand the motivations for a character in a novel read in English class can also be applied to understanding political dynamics in a social studies class.

Note, that a systems-oriented curricula is not a one-size-fits-all structure. The contention is not that systems concepts will provide one complete understanding of either the character in the novel covered in English class or the political dynamics covered in social studies. The value of systems-oriented instruction is that it provides a starting point for understanding new information and can provide insights into a system that conventional curricula often do not provide. I will discuss these insights in the following section. For now let us focus on the transfer of learning.

Research in the transfer of learning has a dubious beginning, as early studies indicating that transfer of learning does occur were used to support the concept of faculty psychology—discussed in Chapter 2—in which the brain is viewed as analogous to a muscle. In this system students were given Classic poems or even lists of numbers to memorize in order to exercise their brains, with the belief that such exercise would transfer to any other mental activity required in the future. Edward L. Thorndike challenged this belief, conducting several studies that fail to show transfer of learning and concluding, “There is no inner necessity for improvement of one function to improve others closely similar to it, due to a subtle transfer of practice effect” (Thorndike & Woodworth 1901, p. 386, qtd in Barnett and Ceci 2002, p. 612).[7]

But while we might easily dismiss the extreme view of transfer suggested by faculty psychology, the idea of transferring learning still lies at the heart of our educational system. At a minimum, we would like to think students will transfer lessons in arithmetic to balance a checkbook or add up a grocery bill and transfer reading and writing lessons to reading a newspaper or filling out a job application. And more ambitious goals regarding critical thinking rest entirely upon the assumption of transfer.

Opinions regarding the prevalence of transfer of learning continue to vary widely. However, Barnett and Ceci (2002) argue that much of the disagreement can be attributed to inconsistencies in experimental designs, and they list several variables that differ between studies, including the nature of the content to be transferred and how the content is initially learned. For example, Barnett and Ceci (2002) distinguish between the transfer of specific facts or rules and that of generalized concepts, suggesting that "general heuristics and principles may transfer more readily than more specific learning" (p. 632). This finding makes sense in that transfer inherently involves some form of generalization. When students are given only specific facts, they must identify general rules on their own before they can transfer what they have learned. When students are taught the general concepts explicitly, they are already one step along toward the transfer of learning. Perhaps even more significantly, Barnett and Ceci (2002) emphasize the importance of "understanding at a deep level," which for Barnett and Cedi means understanding of the material at a "causal" or "structural" level (p. 616). In this context, systems-oriented instruction seems well suited to facilitate transfer of learning across fields of study.

Providing students with cognitive tools that emphasize connections between fields of study has a general pedagogical value inasmuch as it can make learning more meaningful and clarify the applicability of the course material outside of the classroom. But in the context of environmental literacy, understanding the connections between fields of study has an even greater value, since environmental challenges are inherently interdisciplinary. Therefore, understanding a complex environmental issue depends upon students making connections between social, economic, and ecological processes.

In addition to these lateral connections, systems concepts can foster longitudinal connections as well. Since systems concepts and archetypes can be applied across a wide variety of contexts, as students mature they are able to refer to the same familiar concepts to understand more advanced lessons. Returning to our Escalation examples, explication of *The Butter Battle Book* and a discussion of an environmental “race to the bottom”—in which local governments adopt weaker and weaker environmental regulations in order to attract industry—are lessons for two very different audiences. Still, those students who have become familiar with Escalation over the course of their education are likely better prepared to understand its more advanced examples.

Again, these connections are particularly important in the context of environmental education, which has been criticized for its “activity guide mentality” (Knapp 2000; Weilbacher 1997). Knapp (2000) suggests that environmental education often comes in the form of discrete packets rather than a curriculum that builds upon itself from year to year. These “short and sweet strategies,” Knapp argues, “have negated the use of more substantial models that encourage long-term issue investment and, most important, long-term thinking and responsible citizenship behavior on the part of students” (2000, p. 34). Systems concepts offer a basis for connecting disparate lessons into a meaningful whole. For example, a lesson on the water cycle, a staple of environmental education in primary grades, can be followed in later years by a lesson on the carbon cycle to discuss greenhouse gases. While understanding the carbon cycle requires more sophistication (e.g. a basic knowledge of chemistry), both lessons would focus on the various stocks and flows and would emphasize the circular flow of materials. In short, the broad applicability of systems concepts can provide the firm conceptual ground that environmental education lacks without compromising versatility.

3.8 Why Is Learning Systems Concepts Important?

If complex systems are as ubiquitous as I have suggested, then one might reasonably wonder, What's the problem? We have, after all, survived as a species among this complexity for many thousands of years. Presumably, we already possess the insights we need to succeed within such complexity. The answer lies both in the level of complexity in our society and in the spatial and temporal scales at which we must make decisions. In the context of the first point, Joseph Tainter explains, "The citizens of modern complex societies usually do not realize that we are an anomaly of history. Throughout the several million years that recognizable humans are known to have lived, the common political unit was the small, autonomous community, acting independently, and largely self-sufficient" (1978, p. 219).[8] Thus, society has not only become more complex relative to a century ago as illustrated in Chapter 2; it has become uniquely complex on an evolutionary scale.

But it is the second point that plays more of a role in the context of curriculum design. Humans have proven themselves to be able learners. One need only watch an eleven-year-old with an X-box to become convinced of humans' ability to—with practice—master the manipulation of several rapidly changing variables simultaneously. But our day-to-day learning generally takes place in the context of narrow spatial and temporal scales. Humans are not hard-wired to observe or to learn from feedback in complex systems—at least over relatively broad scales—because in systems defined by nonlinear causality, cause-effect relationships are often hidden from the casual observer. There are three fundamental reasons for this. I will describe each of these in turn.

We are already familiar with the first reason: the subtlety of nonlinear causality as explained in Section 3 of this chapter. Identifying causal connections within a causal web can be difficult, particularly when many processes are occurring at the same time. For example, it can be difficult to identify whether the decline of a particular species is a factor of hunting pressure, habitat loss, decline of prey, a natural cycle not previously observed, or some combination of these factors.

The second fundamental difference—time delays—adds yet another level of difficulty in identifying causal links. The causal loops that abound in complex systems often result in delays between causes and their effects. For example, even if humans stopped burning fossil fuels today, global temperature would likely continue to rise for several decades. A casual observer of this phenomenon might conclude that human fossil fuel consumption is not causally connected to global temperatures. But approaching the problem from a systems perspective, one can see that the dynamics of the climate system are such that, although atmospheric levels of carbon dioxide would start declining immediately, the effect of that decline in terms of global temperature will take some time to materialize. A child learns after the first infraction not to touch a hot stove. The lesson would likely not be as clear if the burn and accompanying pain did not manifest until days, weeks, or years after the touching of the stove.

And finally, complex systems exhibit nonlinear behavior that can make casual predictions unreliable. The clearest example of this is exponential growth, the result of a positive feedback loop, capitalized on by the small boy in the following story. The boy, according to most versions of the story, was an acrobat who had just finished a performance for the king. The king enjoyed the performance so much that he asked the boy to stay and perform for the entire month, offering to pay one thousand dollars for each night the boy performed. The boy was poor, and this was more money than he would usually make in several years, but he was also clever. He agreed to perform for the month, but instead of the price that the king offered, the boy asked only one penny the first day, with the pay doubling each day so that on the subsequent days the boy would make two cents, four cents, eight cents, and so on for thirty days. The king, who was not a systems thinker, saw that instead of paying thousands, he would only be paying pennies for the same performances and quickly agreed to the deal. When, after thirty days, they added up the bill, the king found himself bankrupt. Following through with the calculation, you can see that by the end of the thirty days the boy was owed almost eleven million dollars.

Many people are surprised by population growth in the same way that the king was surprised by his bill. For example, a growth rate of 3% or even 5% annually for a country may not sound remarkable. However, at 3% annual growth, a country's population will more than double in 25 years; at 5% it will more than triple in that time. Figure 3-12 shows how the population will increase more and more quickly over the next 50 years even at a constant growth rate. A population of 1,000 will reach almost 12,000 in fifty years with 5% annual growth.

As we have seen, the positive feedback responsible for exponential growth is a basic systems concept, but without an understanding of the concept the behavior will likely be surprising. The literature is filled with situations in business management (e.g. Senge et al. 1999), urban dynamics (e.g. Forrester 1969), and natural resource management (e.g. Gunderson and Holling 2001) where similar surprises have occurred largely due to a lack of understanding of the concepts explained above.

3.9 Conclusion

Thus, in addition to facilitating both lateral and longitudinal connections across a curriculum, an understanding of systems concepts can provide insight into the larger social and ecological systems, which we affect and on which we depend. But in the previous section, I discussed only those difficulties to learning in complex systems that occur when one is actively trying to learn in and about those systems. Arguably, this is not the case, and systems-oriented instruction may be helpful in overcoming a more fundamental hurdle. In the context of education, our problem is not that we do a poor job of teaching our students to deal with complexity; it is that we do not even attempt to address complexity. Our general response to complexity is to ignore it or assume it away, because it can be messy, inconvenient, and confusing. As a result, we continue to be surprised, baffled, and frustrated by the behavior of complex systems.

When one assumes only linear behavior in a system, the solutions to problems may present themselves more clearly and with less effort. But such solutions often miss the mark in significant ways. We are like a sport fisherman who casts his line into the hotel swimming pool—because its water is calm and clear—and becomes disgruntled with his lack of bites. In the following chapter we will look at examples of the kinds of common mistakes and misconceptions that stem from our failure to account for the nonlinear behavior of complex systems.

Notes

[1] Recall from Chapter 1 that Weiner's field of cybernetics focuses on goal-directed systems.

[2] I am using the term "choice" here in the broadest context that need not imply self-awareness. For example, scallops that show a statistical preference for one type of habitat over another may be said to have chosen that habitat.

[3] In large systems, amplification of the effects of stochastic events can also make the behavior of a system inherently unpredictable. This idea is expressed in terms of the butterfly effect, where the effect of air movement by a butterfly's wings can be amplified such that it results in significant changes in weather patterns on the other side of the globe.

[4] Scientists deal with this difficulty using iterative numerical methods to approximate the movement of the bodies involved.

[5] Recently more attention has been given to reinforcing feedback loops that lead to desirable outcomes. For example, the Cold War loop works just as well in the opposite direction. A decrease in U.S. arms would create a decrease in the Soviet's perceived threat, leading to a decrease in Soviet arms, and on through the cycle. The term "virtuous cycles" has been coined to bring attention to how positive feedback loops can be used to achieve desirable outcomes.

[6] See for example Jung (1959).

[7] Recall from Chapter 2 that Thorndike discredited faculty psychology, thereby aiding the shift away from the Classics-based curricula.

[8] Joel de Rosnay makes a similar point: "Today the laws of the economy and ecology come face to face in a type of organization that is new in the history of the ecosystem. It is the nerve center of an immense network of exchanged and consumptions, one of the most complex organizational forms in the social fabric: the city" (1975, p. 27).

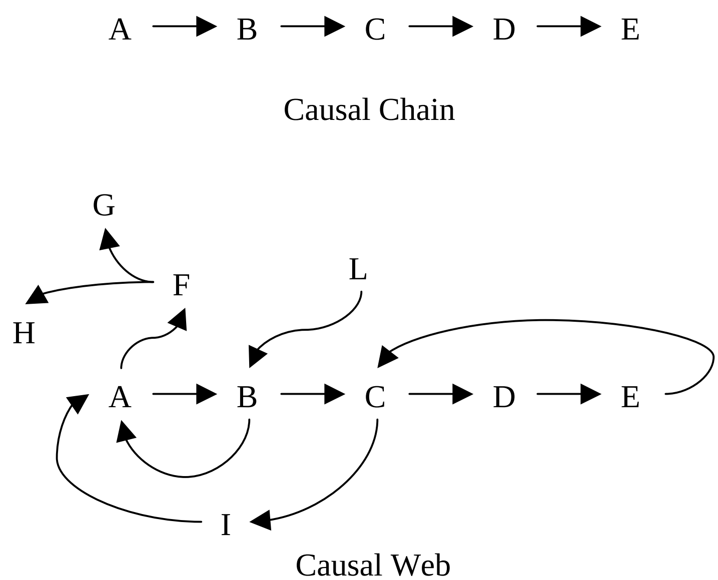


Figure 3-1. Two models for causality—causal chain versus a causal web

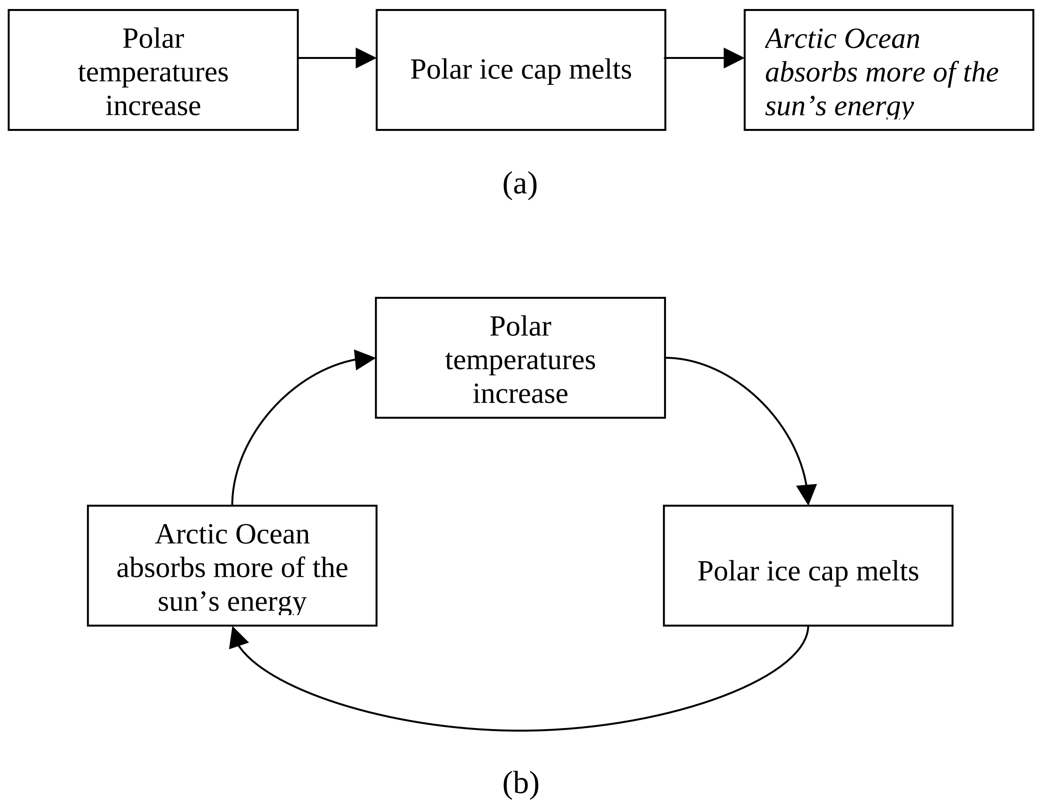


Figure 3-2. Gore's causal connections

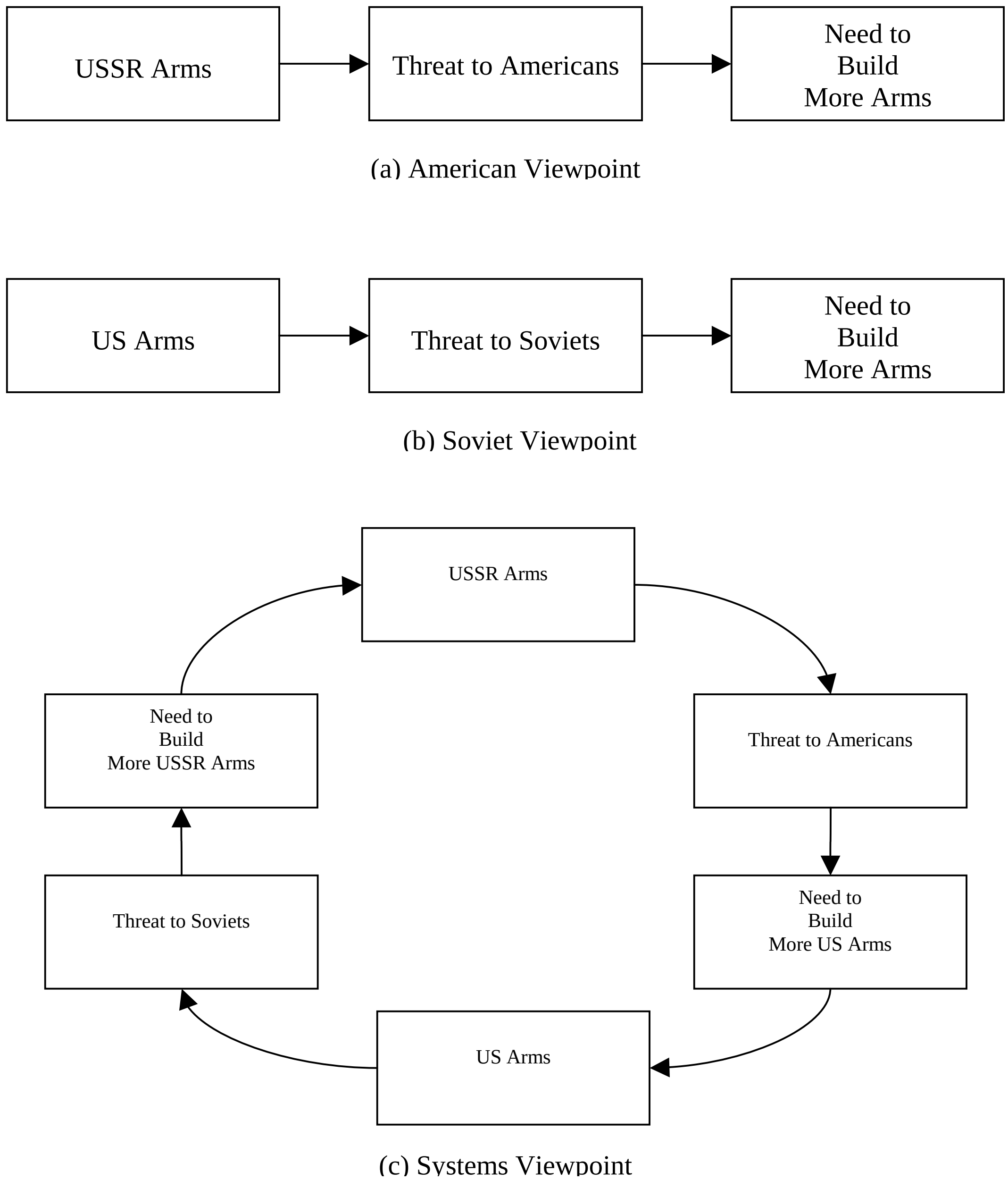


Figure 3-3. Cold war perspectives of the arms race

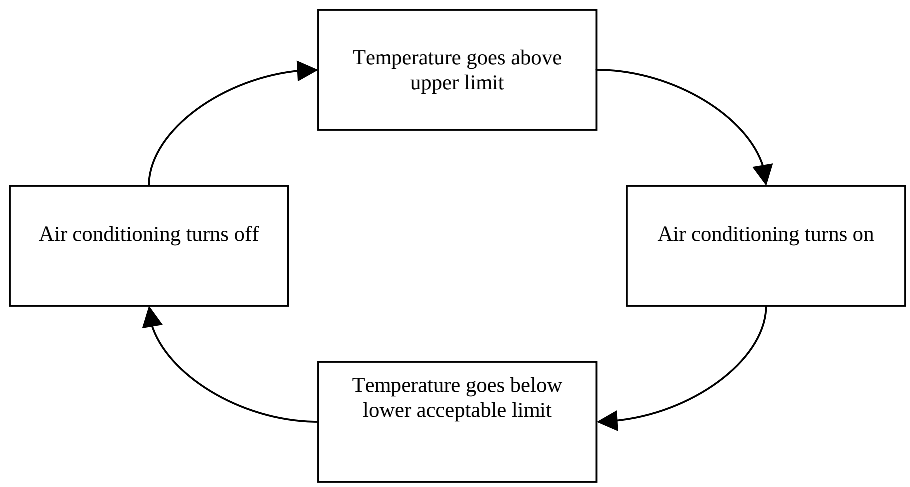


Figure 3-4. Household thermostat as an example of a negative (balancing) feedback loop

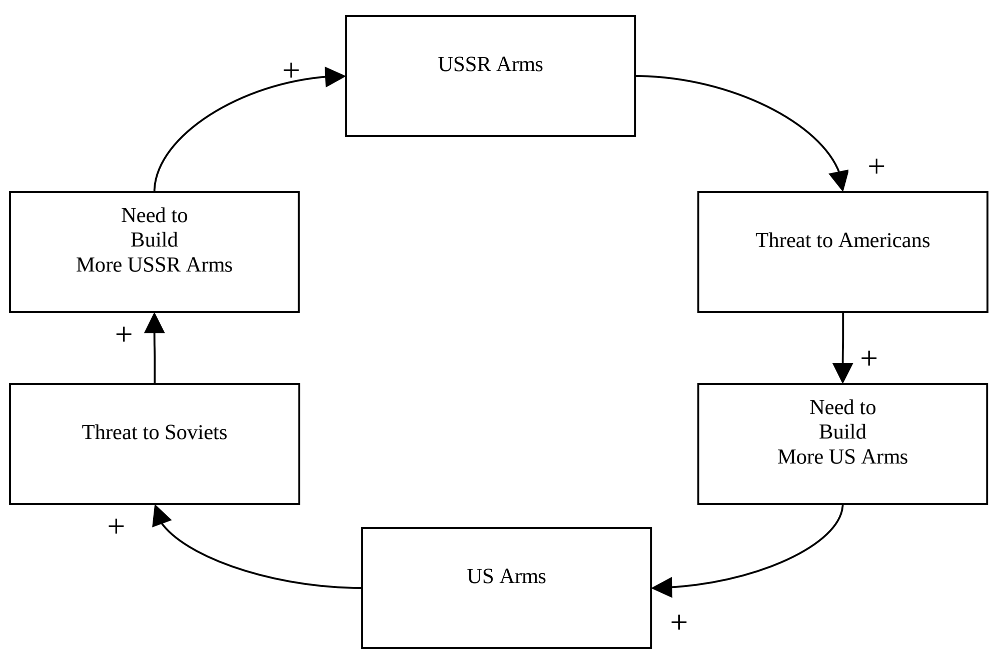


Figure 3-5. Cold war arms race expressed with arrows identified

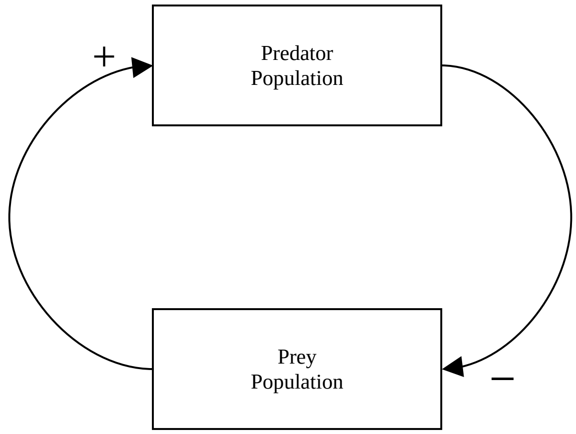


Figure 3-6. Diagram of a system with one prey species and one predator species

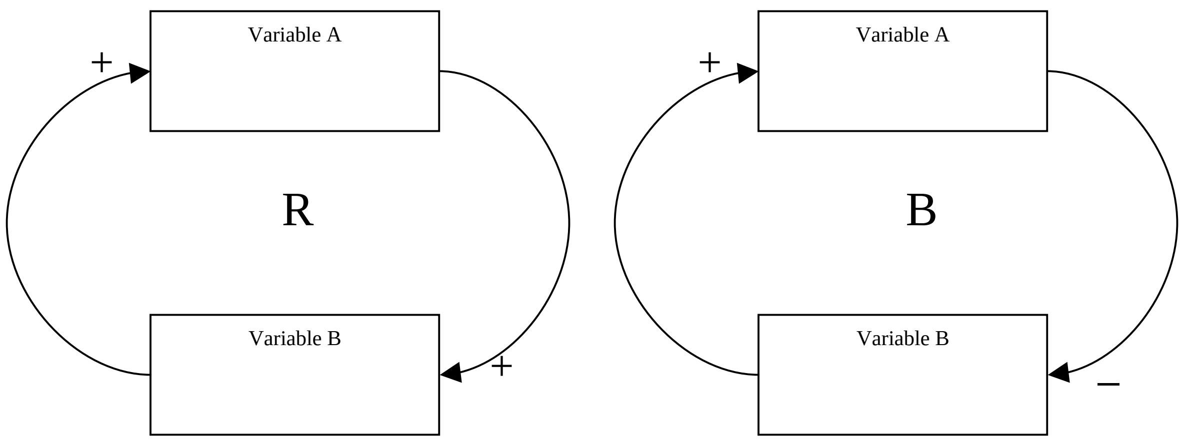
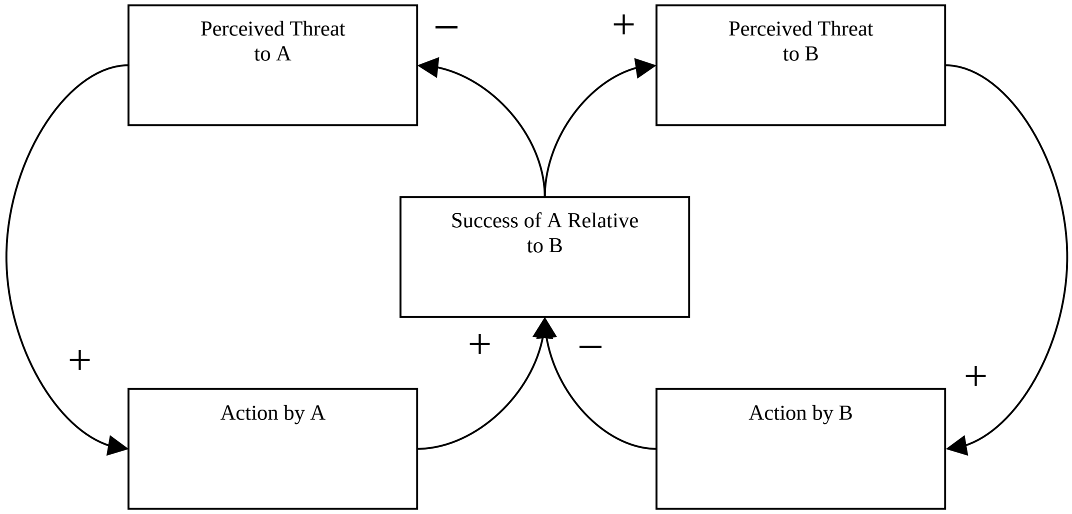
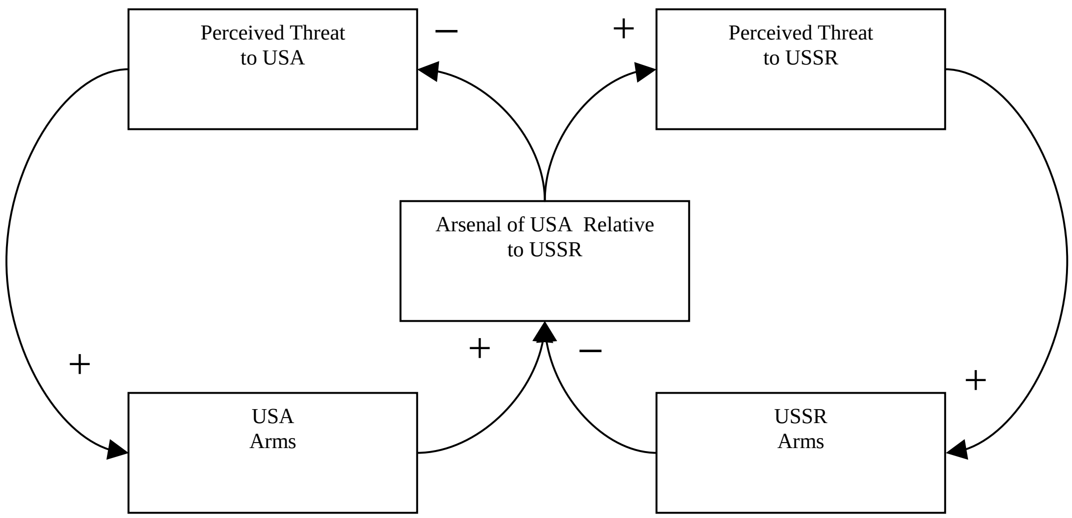


Figure 3-7. Generalized diagrams for reinforcing (left) and balancing (right) loops

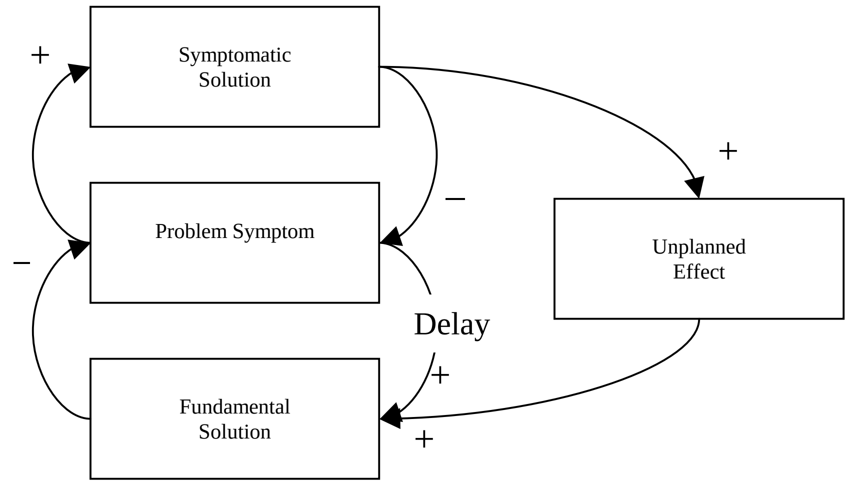


(a) Generic Diagram

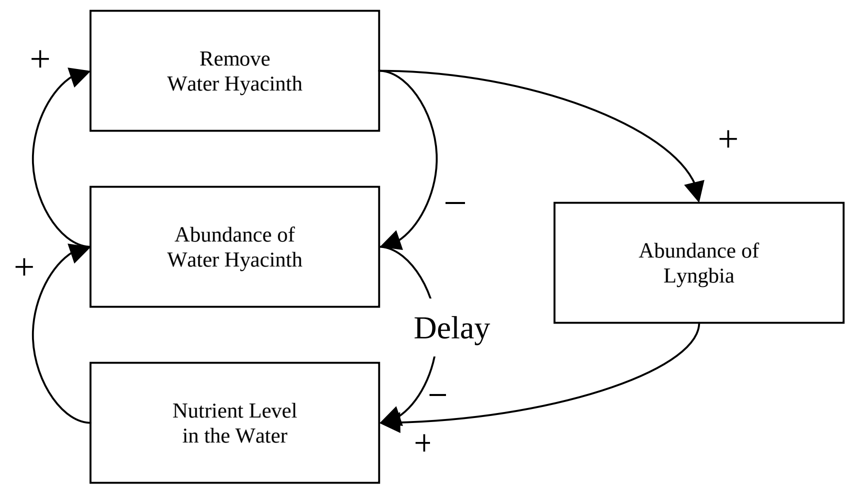


(b) Cold War Diagram

Figure 3-8. Escalation archetype

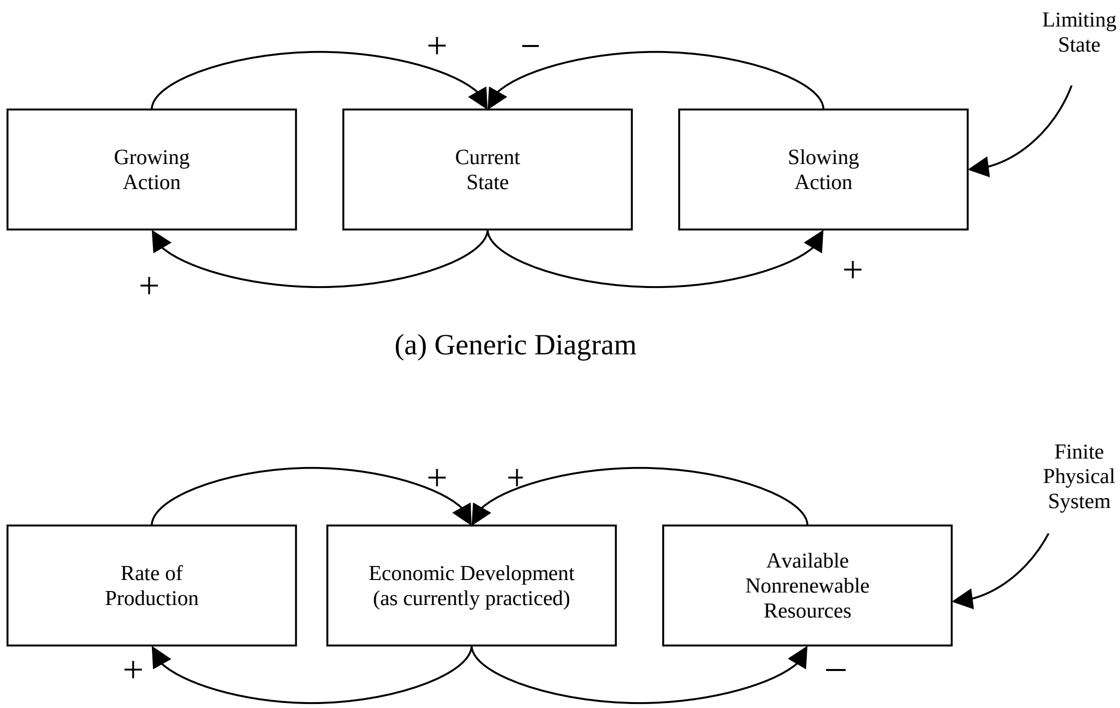


(a) Generic Diagram



(b) Spring Management

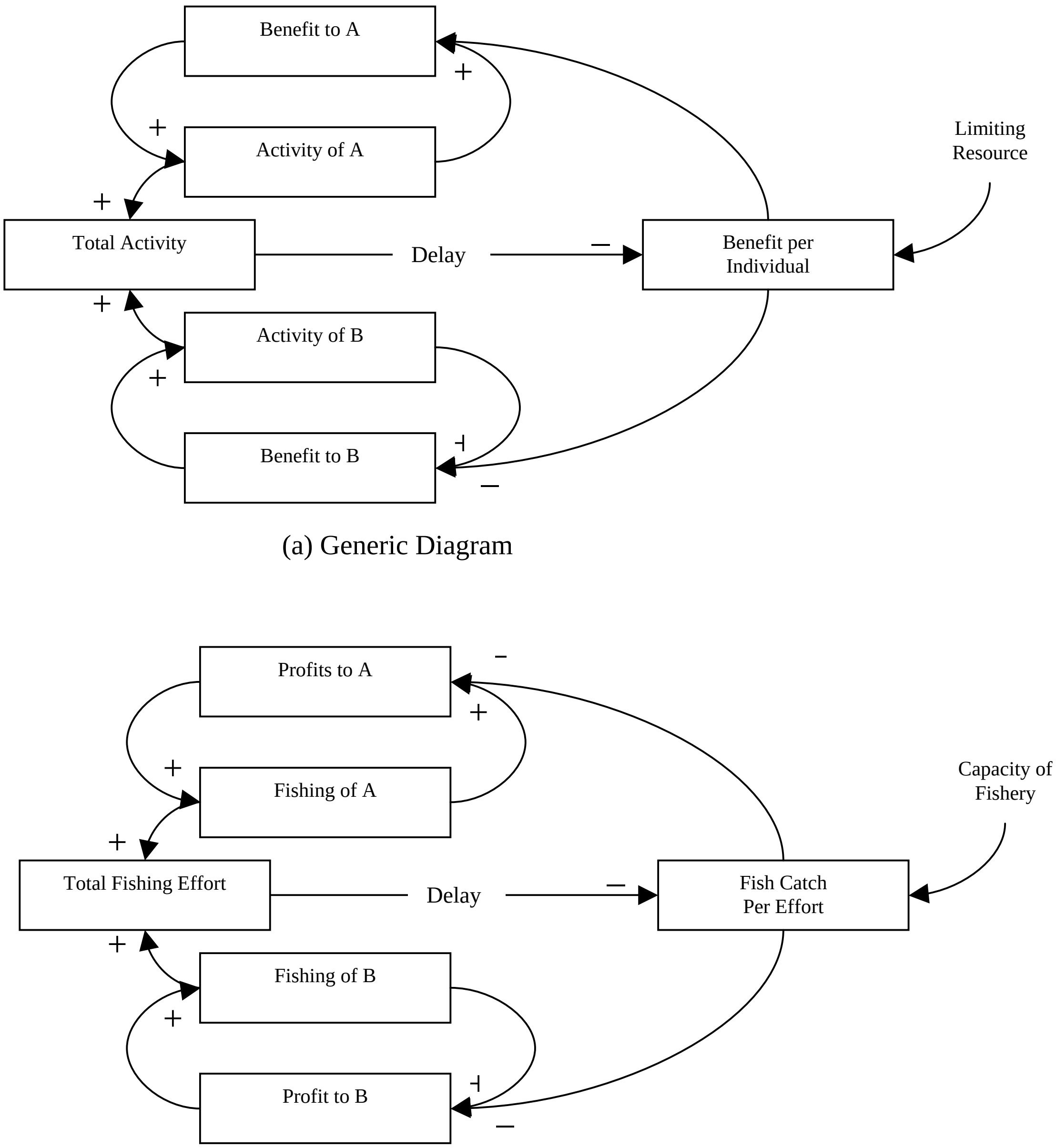
Figure 3-9. Shifting the burden



(a) Generic Diagram

(b) Economic Development

Figure 3-10. Limits to growth



(a) Generic Diagram

(b) Economic Development

Figure 3-11. Tragedy of the commons

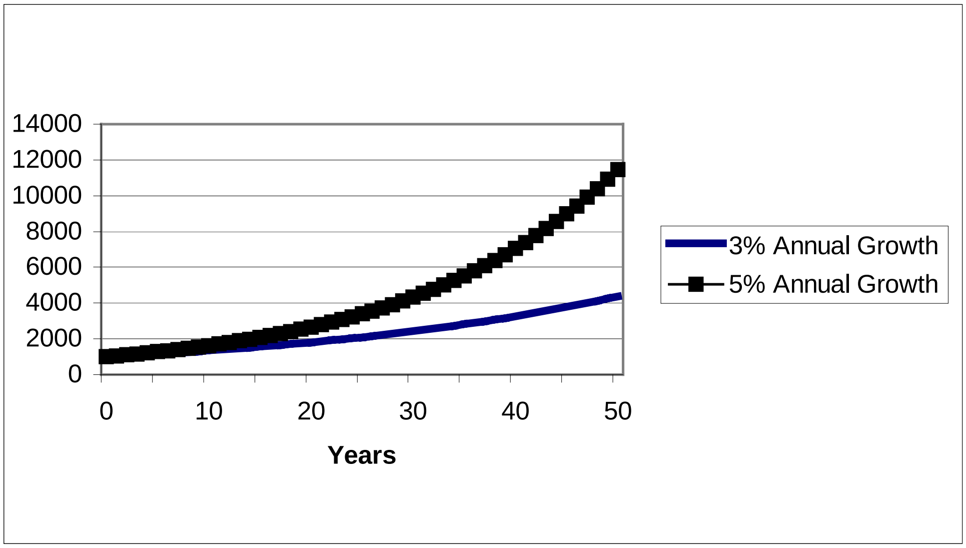


Figure 3-12. Exponential population growth

CHAPTER 4
SYSTEMS THINKING IN THE ROLE OF ENVIRONMENTAL DECISION-MAKING

4.1 Introduction

Armed with a conceptual understanding of systems thinking and systems-oriented instruction, the discerning reader may be tempted to ask, So what? What difference does systems thinking make on a practical level? In this chapter I address this question in the context of decision-making. Perception is an active process, involving both filtering and interpreting stimuli received from sensory organs. Therefore, what we learn through reading, watching, listening, and experiencing depends largely on the mental models we use to interpret our sense experience. In this chapter, I review a number of studies that provide insight into how our mental models affect both what we learn and the decisions that we make.

Leigh Thompson (2001) titles the appendix to her text on negotiation with the question, "Are you a rational person?" In the appendix she identifies the advantages of rational thinking and describes the multiattribute utility technique (MAUT) for making rational decisions. MAUT consists of five main tasks: "(1) identify the alternatives, (2) identify dimensions or attributes of the alternatives, (3) evaluate the utility associated with each dimension, (4) weight or prioritize each dimension in terms of importance, and (5) make a final choice" (2001, p. 295). To illustrate the technique Thompson takes the reader through a choice between two different graduate programs, assigning a numerical weight to each key attribute (e.g., tuition cost, reputation, climate, and culture), assigning each school a score for each attribute, and calculating a winner.

The process seems straightforward enough, but, Thompson concedes, "Sometimes we must make decisions when the alternatives are uncertain or unknown" (2001, p. 297). Most of the researchers cited in this chapter would make this point much more strongly, suggesting that most of the decisions people make involve alternatives that are uncertain or unknown. In fact, MAUT sounds very much like the economic characterization of rational choice that Herbert Simon (1956) criticized several decades ago, arguing that such rational choice processes simply require more information and computational ability than people generally have. He suggested that to make the concept of rational choice decision-making useful, one must account for “the limitations upon the capacities and complexity of the organism” making the decision (1956).

Psychologist Robin Hogarth (1980, pp. 4-6) describes these limitations with respect to humans. First, human perception is selective rather than comprehensive. Secondly, people are limited in the amount of information they can take in at one time. Consequently, “processing is mainly done in a sequential manner [which] can be misleading in the sense that the actual sequence in which information is processed may be biased by a person’s judgment” (Hogarth 1980, p. 4). Thirdly, people have a limited capacity for processing information. When provided with an overabundance of information, much of it becomes aggregated or lost. And finally, people have limited memory. Therefore, rather than accessing full memory like a computer, we reconstruct memories based on sparse facts. In short, humans do not perceive completely, process information quickly or efficiently, or retain information effectively.

Much of the research regarding decision-making revolves around how people deal with these deficiencies. Hogarth concedes, “From the above, one might conclude that people…can be likened to ineffective computers” (1980, p. 6). But he dismisses this view because “the computer analogy is inappropriate” (p. 6). Humans have imagination, emotions, and creativity “that would be extraordinarily difficult to program into computers. Furthermore, people attach *meaning* to information and such meaning is often the clue to understanding how human thought processes work” (p. 6) [emphasis in text]. For Hogarth, people derive this meaning from their “intuitive understanding of how events are related" (1980, p. 42). But this necessary reliance on intuition and meaning can create problems.

Amos Tversky and Daniel Kahneman made careers out of illustrating the chronic errors people make as a result of irrational biases (See Kahneman and Tversky 1974, 1981, 1996, and Kahneman *et al*. 1982). They suggest that "people rely on a limited number of heuristic principles which reduce complex tasks of assessing probabilities and predicting values to simpler judgmental operations. In general these heuristics are quite useful, but sometimes they lead to severe and systematic errors" (Kahneman and Tversky 1974, p. 1124). While Kahneman and Tversky often emphasize these severe and systematic errors, Gerd Gigerenzer (See Gigerenzer 1996 and Gigerenzer *et al*. 1999) takes a more positive view of these heuristic principles, suggesting that the heuristics prevailing in decision-making have evolved as a tool for dealing efficiently with unmanageable amounts of information. In the following section I will review this debate between Kahneman and Tversky and Gigerenzer in more detail both because it involves key figures in behavioral decision research and because it provides a structure for introducing the concepts that are central to much of the research within this field.

Recommendations for improving judgment revolve around improving individuals' mental models of the situation in question. Therefore, in the third section I will introduce the concept of mental models and discuss their role in decision-making. In the fourth section, I will continue the discussion of heuristics in the context of complex system dynamics. Systems theorists have identified biases analogous to those studied by Kahneman and Tversky, among others. In this context, systems theorists generally suggest improving judgment through a shift toward systems thinking (i.e., a more dynamic mental model). In the final section, I will make some suggestions regarding the role of systems-oriented education in affecting environmental judgment and behavior.

4.2 Simple Heuristics

Kahneman and Tversky (1974) have described a wide array of heuristics and the biases to which they lead. They divide these heuristics into three main categories. The first is Representativeness, which refers to an individual's tendency to base a prediction about an event not on statistics or probability, but on the degree to which the event correlates (or represents) preconceived notions. For example, subjects were given a description of a person said to have been picked randomly from a group of 70 engineers and 30 lawyers. When asked for the probability that the man was a lawyer, subjects based their answers on how well the description fit a stereotypical engineer or lawyer despite the clear numerical cues. Moreover, the verbal description affected subjects' responses even when it contained no relevant information. For example, when no description was given, participants answered according to the numerical cues (70% chance of being an engineer). But the following description changed the response from participants: "Dick is a 30 year old man. He is married with no children. A man of high ability and high motivation, he promises to be quite successful in his field. He is well liked by his colleagues" (Kahneman and Tversky 1974, p. 1125). Armed with this information, which provides no clues to the question at hand, participants disregarded the numerical cues and claimed there was an equal chance of it being a lawyer or an engineer, leading Kahneman and Tversky to conclude that "people respond differently when given no evidence than when given worthless evidence" (1974, p. 1125).

Kahneman and Tversky's second category, Availability, refers to individuals' tendency to estimate the frequency of an event based on how many instances of such an event come easily to mind. For example, most study participants believe that there are more words that start with *r* than there are words that have *r* as the third letter, simply because it is easier to think of words that start with *r*. And the final category, Adjustment and Anchoring, refers to a tendency to allow one's first estimate of a value to anchor subsequent estimates. For example, when individuals were asked to estimate the percentage of African countries in the United Nations, those who were led to start from a very low first estimate averaged much lower guesses that those led to start from a very high estimate. The initial values given to the participants anchored their subsequent guesses.

Put simply, Kahneman and Tversky argue that when relying on their intuitive judgments, people often make the same kind of mistakes. Furthermore, these mistakes cannot be attributed to apathy or misconceptions of the uneducated. Biases persisted despite incentives and over subject groups representing varied education (1974, 1981). But Kahneman and Tversky have been criticized for what some see as an overly negative view, basically identifying humans as chronically poor decision-makers. For example, L.J. Cohen (1981) and Lola L. Lopes (1991) both criticize Kahneman and Tversky for structuring their experiments such that their subjects look particularly irrational.[1] Without getting into the details of these criticisms, we might do well to pay attention to an early observation Masanao Toda (1962) made regarding the rather contrived nature of many judgment experiments. "It is needless," he suggested, "to point out how difficult it is for us to make threshold judgments, to not be tricked by an optical illusion, to remember nonsense syllables." He conceded that humans are "incredibly stupid in an experimental room. On the other hand…man drives a car, plays complicated games, designs computers, and organizes society." In light of these accomplishments, suggested Toda, one studying human judgment might start with the assumption that humans are efficient problem solvers.

Gigerenzer and his associates seem to agree with Toda's assumption. Their system of fast and frugal heuristics can be seen as a partial explanation for how humans have become such efficient problem solvers. Gigerenzer suggests, along with Kahneman and Tversky, that people do not make their choices based purely on reason and mathematical calculations. But for Gigerenzer, the simplifying heuristics invoke admiration rather than concern. Where Kahneman and Tversky (1974) criticize individuals for making judgments incompatible "with the entire web of beliefs held by the individual," Gigerenzer (1999) marvels at their relatively high rates of success with extremely limited time and information. For Gigerenzer real-world results take precedence over abstract logical coherence. He explains,

> If you believe that there is a probability of 90% that Elvis is still alive and a probability of 10% that he is not, your beliefs are at least coherent, in that you give the two opposite possibilities together a 100% chance, as probability theory says you must. But if they lead you to spend long hours in cornfields waiting for his UFO to land, these beliefs are not doing you much real-world good. (Gigerenzer and Todd 1999, p. 21)

In this context, he argues that concepts like unbounded rationality and optimization have no place in discussion of real-life decision-making, simply because they require far too much information and computation time to be practical. Instead, he suggests the concept of ecological rationality, where the success of a heuristic is measured by "its performance with the actual requirements of its environment, which can include making accurate decisions, in a minimal amount of time, and using a minimal amount of information" (Gigerenzer and Todd. 1999, p. 22).

The result of this shift in focus is arguably a much more positive outlook on people's ability for sound judgment. Gigerenzer explains, "whereas [Kahneman's and Tversky's] heuristics-and-biases program portrays heuristics as a frequent hindrance to sound reasoning, rendering *Homo sapiens* not so sapient, we see fast and frugal heuristics as enabling us to make reasonable decisions and behave adaptively in our environment—*Homo sapiens* would be lost without them" (Gigerenzer and Todd 1999, p. 29). And indeed, in chapters with titles like "How Ignorance Makes Us Smart," Gigerenzer and his group provide some compelling evidence for the success of their fast and frugal heuristics (in terms of accuracy and/or computation time) over far more sophisticated models.

For example, Gigerenzer *et al*. make the distinction between completely unrecognized objects, recognized objects, and objects about which something is known beyond recognition. For situations in which one's knowledge is extremely limited, suggest Gigerenzer and Todd (1999), the distinction between recognized and unrecognized objects can be a useful heuristic even when nothing else is known. For instance, rats have been shown to prefer food that they have eaten before or smelled on another rat's breath to completely foreign food, reducing the likelihood of ingesting something poisonous. In fact the heuristic of simply choosing a recognizable object over an unrecognizable one is most effective when nothing else is known. Gigerenzer and Todd (1999, p. 44) show that Germans predict the relative size of U.S. cities more accurately than Americans, not because they know more about the U.S., but because they know less. The Germans simply chose the cities they recognized. They accurately identified San Diego as bigger than San Antonio largely because they had heard of San Diego, but had not heard of San Antonio. The Americans, recognizing both cities, often made the problem much more complicated and were significantly less successful.

When an individual does know more about a situation than mere recognition, she can use the "take the best" heuristic in which only one factor is used to distinguish between two competing objects or options, rather than trying to weigh an entire list of pros and cons for each option (Gigerenzer and Todd 1999, p. 75). Gigerenzer's research group provides numerous examples of studies in which these simple methods competed well against—or even outperformed—more sophisticated decision-making tools. But all of these studies are set up such that the participant is either choosing between two objects or choosing a set of objects from a list. That is, there is a finite and given list of choices from which the participant or computer model must pick. Berndt Brehmer (1980) suggests that studies like these do not say anything about how an individual makes a decision when the problem is not so structured. That is, they say nothing about how an individual develops her own list of possible courses of action or about how she judges some information as more relevant than the rest.

In this sense the suggestion that the structure of Kahneman's and Tversky's studies lead to a skewed and incomplete picture of heuristics can be directed at Gigerenzer as well. What we have are two sets of experiments showing that in some instances heuristics lead to erroneous biases and in others they perform quite effectively.[2] While Gigerenzer (1991) makes some direct criticisms of Kahneman's and Tversky's methods, most of the fast and frugal heuristics documented by Gigerenzer do not directly contradict the biases identified by Kahneman and Tversky. The point here is not to prove that heuristics are generally good or generally bad. A more useful direction would be to distinguish between situations where heuristics lead to sound (or at least ecologically rational) judgments and when they lead to erroneous and deleterious biases. In section IV, I shift the discussion from probability and simple recognition to predictions in the context of complex systems. As we shall see, common biases in this area may be particularly important with respect to environmental judgment. However, a brief aside into mental models may be necessary in order to better understand that discussion. This aside comprises the following section.

4.3 Framing Problems: The Use of Mental Models

Out of all the misleading biases they have documented, Kahneman and Tversky (1974) identify, as particularly confusing, individuals' failure to recognize statistical phenomena that they see regularly. "What is perhaps surprising," they explain, "is the failure of people to infer from lifelong experience such fundamental statistical rules as regression toward the mean, or the effect of sample size on sampling variability. Although everyone is exposed, in the normal course of life, to numerous examples from which these rules could have been induced, very few people discover the principles of sampling and regression on their own" (1974). Their explanation for this failure is that "relevant instances are not coded appropriately" (1974). In other words, although the information necessary for apprehending such statistical phenomena is available, people do not organize the information in a way amenable to statistical observations.

This theme of problem coding or framing is common among researchers in decision-making. Hillel Einborn and Robin Hogarth (1978) suggest that not only do people have biased judgments, but they also have a high level of confidence in those judgments.[3] They attribute the persistence of biases like those identified by Kahneman and Tversky partially to "how outcomes are coded and interpreted" (Einborn and Hogarth 1978). In this case the context for errors is not just a lack of statistical understanding, but an ignorance of formal reasoning. Einborn and Hogarth (1978) cite a series of experiments by Peter Wason as well as their own experiments, showing that people trying to identify a rule will use confirming evidence, rather than more useful disconfirming evidence. For example, participants were given a series of three numbers (e.g., 2, 4, 6) and asked to identify the rule that the series was following. To solve the problem the participants could come up with their own series and ask the experimenter if their series followed the rule. In this experiment, participants used confirming evidence (e.g., testing out an even-number hypothesis by suggesting another series with even numbers) rather than disconfirming evidence (e.g., putting an odd number in to see if it failed the rule as expected).

Einborn and Hogarth (1978) conclude that people simply lack the formal reasoning concepts necessary for coding information in a more logical way.

Brehmer (1980) comes to a similar conclusion regarding individuals' tendency to use representation and availability to make estimates rather than following the rules of probability. He explains that "subjects do not seem to be able to learn to perform optimally in probabilistic inference tasks… due to a lack of schemata for handling the probabilistic aspect of the world….In short probability must be invented [or learned] before it can be detected" (1980). For Brehmer, one is unlikely to see probability rules through experience unless one has been taught to observe phenomena in the context of probability. That is, one's mental model must include the concept of probability.[4]

Jill Larkin (1983) discusses a similar problem in the context of physics. She asks, "Inasmuch as people are good at predicting the outcome of physical interactions in the world around them, why are they so bad at physics, even the branch of physics (mechanics) that deals with the interaction of everyday objects" (1983). Her answer echoes those offered by the other authors cited in this section, but Larkin comes at the problem from the opposite side. Whereas Kahneman, Tversky, and Brehmer are interested in why people do not know basic rules of statistics and probability simply from experience, Larkin focuses on how best to use student experience to give students a basic understanding of physics concepts.

Thus, where Kahneman, Tversky, and Brehmer suggest that one can better learn from experiences by coding those experiences in the context of abstract statistics and probability rules, Larkin (1983) suggests that abstract physics concepts can be more effectively learned if coded to fit students' experiences. To test her hypothesis, Larkin divided a class of beginning physics students into two groups. Both groups received a basic introduction to seven principles describing DC circuits. The first group learned how to derive equations from these principles and practiced how to manipulate those equations algebraically. The second group continued their training by learning the concepts in the context of physical applications (e.g., by analogies to fluid flow). At the end of the training the second group was significantly more successful in solving complicated physics problems that required using a combination of several of the concepts learned at the beginning of the class.

The effectiveness of relating the concepts to a directly observable situation over practicing algebraic manipulation of abstract equations may not be surprising, but Larkin goes further, asserting that the ability to visualize abstract concepts in physical terms is what separates the skilled scientist from the novice. She states, “[T]he process of mentally simulating events so as to predict their outcome, a facility possessed by most people for common contexts, is extended and refined in a skilled scientist to become a sharp and crucial intuition that can be used in solving difficult, complex or extraordinary problems” (1983). But Gentner and Gentner (1983) suggest that learning concepts by analogy to directly observable systems can lead to their own systemic biases. Still in the context of DC circuits, Gentner and Gentner study how the choice of analogy—in this case fluid flowing through pipes or a crowd of people flowing through a narrow hallway—affected participants’ answers to circuitry problems. Before the experiment, Gentner and Gentner were able to accurately predict where participants would answer incorrectly based on the analogy a participant was using to understand the problem. They explain that “the use of different analogies leads to systematic differences in the patterns of inferences” (Gentner and Gentner 1983).

Thus, Larkin seems to oversimplify the situation, exaggerating the role of visualization. Were a skillful scientist to rely solely on her ability to visualize an event in terms of an analogous, directly observable system, her science would be only as good as her analogy. But Larkin's ideas are important to this discussion, because most researchers of decision-making are not interested in producing skillful scientists. Kahneman's and Tversky's call for improved statistical judgment does not require a society of skillful statisticians. Rather, it requires a society of people who have a basic understanding of statistics and are able to apply that understanding to enhance their judgments. Gentner and Gentner (1983) effectively show that using fluid flow or people-flow analogies to convey principles of circuitry will not necessarily produce skillful scientists. What it will produce is a group of people with a fundamental understanding of some basic circuitry principles.

If, as many of the authors here suggest, a better understanding of statistics, logic, probability, etc. is necessary to avoid heuristic biases, then establishing the level of desired proficiency in statistics, logic, probability, etc. seems to be an important issue—one that these researchers leave largely untouched. Kahneman and Tversky point out, "A person could conceivably learn whether his judgments are externally calibrated by keeping a tally of the proportion of events that actually occur among those to which he assigns the same probability," but they concede that "it is not natural to group events by their judged probability" (1974). Similarly, Einborn and Hogarth suggest encouraging people to encode events "not by their substantive content but by judged probability" (1978). In addition, they suggest teaching people formal logic so that they will understand the value of seeking disconfirming information (Einborn and Hogarth 1978). R.E. Nisbett et al. (1983) follow the same path, suggesting formal statistical training to improve judgment. Such changes may likely reduce heuristic biases as suggested. However, for these recommendations to have any applicable power, researchers need to identify more specifically what levels of proficiency in these areas are necessary to affect decision-making. Furthermore, in that our preference for verbal content over statistical or probabilistic values seems to be particularly strong, one might be wise to accept this preference as given when designing a program to improve judgment.

It is also worth mentioning that some question the value of more sophisticated mental models altogether. James Doyle (1997) suggests, "Evidence is emerging that more complex cognitive structures, such as mental models of systems, are…not necessarily related to behavior in ways that can easily predicted *a priori*." Doyle makes an important point here. Implicit in the calls for more statistical and probability training is the view of humans primarily as calculating machines. The assumption is that providing individuals with better calculating methods will improve their judgment. The idea has some validity, but it is incomplete.[5] Doyle simply means to challenge this assumption and question the connection between mental models and behavior, but a closer look at the evidence he cites provides a more important lesson and leaves plenty of room for hope in the ability of effecting changes in behavior by changing mental models.

For example, Doyle (1997) cites Willet Kempton's (1986) study on people's perception of heating systems. Kempton (1986) describes two different theories of how home heating systems work—(1) the feedback theory, which is the scientifically accurate one, and (2) the valve theory where the thermostat setting is thought to control the magnitude of the flow of heat through the system. Despite the inaccuracy of the model provided by the valve theory, Kempton suggests that valve theorists may be more likely than feedback theorists to exhibit energy-saving behavior. In essence, Kempton suggests a distinction between accurate mental models and useful mental models.

Donald Norman (1983) makes a similar distinction regarding people's use of calculators. Doyle (1997) paraphrases Norman's findings as follows: "old behavioral habits persisted even after the adoption of improved mental models that acknowledge that the behaviors are unnecessary." The description is not inaccurate, but left as is, it seems to imply a lack of connection between the behavior and mental models of Norman's subjects. Upon reading Norman's article, one finds that this is not the case. In fact, Norman's subjects did not fully adopt the improved mental models. Norman (1983) attributes the discrepancy between behavior and the improved mental models to both a desire for simplification and a lack of confidence in the new mental models. For example, subjects would hit CLEAR several times even when they knew that hitting the button once would suffice. Norman notes that some models of calculators in the study required depressing CLEAR more than once to clear all the registers. He explains, "The rule to hit the CLEAR button excessively allows the user to avoid keeping an accurate count of the operation. Moreover, it provides a rule that is functional on all calculators, regardless of design" (1983). In other words, it is simply easier to hit CLEAR a few extra times than it is to keep close count of how many CLEARs are actually necessary. The subjects simplified the improved mental model into an arguably more robust model. Rather than arguing against the connection between behavior and mental models, Norman (1983) simply provides a list of traits a mental model must have in order to affect behavior: *learnability*, *functionality*, and *usability*.

Thus, we might push Doyle's (1997) claim that "the relationship [between mental models and behavior] is often complex and counterintuitive" even further. The complexity to which Doyle refers can be explained in the context of Norman's three terms. For Norman's subjects, *learnability* was not the problem. They could understand how to operate the calculators efficiently. However, the more complicated models that included details of specific memory functions lacked *usability*. Specific operating instructions were too cumbersome to keep track of, so the more accurate model was discarded for a more usable one. In this context, we can view Kempton's (1986) results to be challenging the *functionality* of the feedback theory of heating systems and suggests that the scientifically inaccurate valve theory is actually more functional (i.e., leads to more desirable behavior).

These studies suggest that while scientists would like their subjects to be able to provide scientifically accurate models of the system at hand, perhaps it would be more profitable to focus more on the potential for a model to be adopted and the type of behavior the mental models elicit. Norman (1983) observes, "People's mental models are apt to be deficient in a number of ways, perhaps including contradictory, erroneous, and unnecessary concepts." The key, analogous to the conflict regarding heuristics, is to identify which of those deficiencies we can live with, and which need to be improved upon.

To summarize, I began this discussion with a review of the relative agreement among researchers of decision-making that people do not behave like purely rational beings. Instead, when faced with uncertainty, individuals use simplifying heuristics to make their decisions. I summarized Kahneman and Tversky's argument that these simplifying heuristics often lead to systematic biases and Gigerenzer *et al*.'s argument that these heuristics are—in real situations—often more successful than sophisticated calculating methods. Glossing over the methodological disagreement between these two groups, I suggested that collectively their work shows that the success of simplifying heuristics varies depending on the context of the problem.

In this section I reviewed several studies regarding what is commonly perceived to be a major factor in heuristic biases: the mental models and concepts used to code information. These studies showed that not only do mental models greatly affect how people interpret—and learn from—their experiences, but previous experience greatly affects how one understands new mental models. I concluded this section by reviewing some of the challenges presented by this research including the development of realistic educational goals regarding mental models. In the following section, I review the study of heuristics and biases in the context of complex systems.

4.4 Decisions in the Context of Complexity

So far, our focus regarding conceptual models has been on logic and mathematics. Generally speaking, when researchers with this focus refer to *complex* problems, they are referring to detailed complexity. Recall from Chapter 2 that problems with high detailed complexity involve a high degree of information and often call for sophisticated calculations. For example, John Payne *et al*. (1992) equate complexity with the number of choices and amount of information available. They explain, “Perhaps the most well-established task-complexity effect is the impact of changes in the number of alternatives available. . . . Varying the amount of attribute information is another way to manipulate decision complexity.” Thus, you achieve complexity by overloading people either with choices or with information.

But there is another area of behavioral decision research, which deals more specifically with dynamic complexity. Recall that dynamic complexity arises from interactions between parts of a system. Jay Forester (1996) suggests that poor judgment in the context of dynamic complexity often results from a failure to understand the nature of complex relationships. He explains, “Most of our intuitive learning comes from very simple systems. The truths learned from simple systems are often completely opposite from the behavior of more complex systems” (1996).

Forrester’s (1996) example of a simple system is filling a glass of water. Here, one has a stock (water in the glass), a flow (water flowing from the faucet), and a feedback (the level of water in the glass that signifies to the individual when to turn off the faucet). He contrasts this from a far more complex model of a nation's population and economy, suggesting that the large number of interrelationships makes an intuitive understanding of the model difficult.[6] While Forrester's examples provide a good basis for distinction between simple and complex systems, one might concede that even those with an understanding of systems concepts might not be able to predict specific behavior of a highly complex system.

John Sterman and Linda Booth Sweeney report a somewhat more troubling ignorance of systems dynamics. They have conducted two studies that illustrate how even highly educated individuals lack the mental models necessary to understand and interpret complex systems on even a general level. In their first study (Booth Sweeney and Sterman 2000) they used a model of a bathtub in order to test participants' knowledge of stocks and flows. In the model there was an inflow (the faucet), a storage (the bathtub), and an outflow (the drain).[7] In order to make the systems as simple as possible, flow through the drain did not depend on water level in the tub.

Even with these simplifications, the participants, consisting of graduate student in MIT's Sloan School of Management, had difficulty graphing the water volume in the bathtub based on given manipulations to the inflow. Booth Sweeney and Sterman (2000) suggest that subjects often "rely on a heuristic that matches the shape of the output of the system to the shape of the input." For example, instantaneous changes in the inflow (say from 25 gallons per minute to 75 gallons per minute) translated for many subjects into instantaneous changes in volume (say, from 100 gallons to 150 gallons).

They note that the "two features that subjects find problematic—the slope of the stock is the net flow, and the change in the stock over an interval is the area enclosed by the net rate in that interval—are the two fundamental concepts of calculus," but explain that training in calculus, which all the participants had at some level, "did not translate into an intuitive appreciation of accumulations, of stocks and flows" (2000). After identifying a host of situations in which distinguishing between stocks and flows would be necessary—including atmospheric CO_2 versus CO_2 emissions and national debt versus the annual deficit—they suggest that pedagogical goals should be focused more on developing basic intuition of systems rather than formal mathematics. Booth Sweeney and Sterman (2000) explain, "[O]ur results suggest that good mathematics training alone is not sufficient to develop a practical, common-sense understanding of the most basic building blocks of complex systems." Similar to Larkin's (1983) suggestion in the context of physics, Booth Sweeney and Sterman emphasize the importance of relating system dynamics to everyday experience in order to develop a more intuitive sense of those dynamics.

In their second study Sterman and Booth Sweeney (2002) look more directly at individuals' perception of global warming. They cite a Center for Policy Attitude poll, showing that while most Americans believe that global warming is real, only a small percentage believe that it is a "serious and pressing problem." In order to explore these perceptions more deeply, Sterman and Booth Sweeney provided participants with "a short description of the climate system graphs showing the history of human CO_2 emissions, the concentration of CO_2 in the atmosphere, and global mean temperature" and asked them "to specify the likely response of the climate" based on various CO_2 emissions scenarios (e.g. emissions holding constant and emissions dropping to zero). No calculations were involved here. The analysis was entirely qualitative. However, participants, consisting this time of graduate students from MIT, Harvard, and the University of Chicago, were found to hold troubling views about global warming, following a heuristic similar to that found in the previous study:

> The vast majority believe that temperature should follow the same pattern as CO_2 concentration, rising when CO_2 is rising and falling when CO_2 is falling…. Similarly, most

> believe that stabilizing CO_2 concentrations can be accomplished by stabilizing emissions near current rates, when emissions must fall significantly, to the removal rate, for concentrations to stabilize. (Sterman and Booth Sweeney 2002)

Sterman and Booth Sweeney suggest that misperceptions regarding stocks and flows—both of CO_2 and of heat—promote complacent attitudes regarding global warming. Generally, participants expressed a wait-and-see approach, in which action to reduce global warming could be taken if and when conditions worsened significantly. Others recommended extremely minor reductions in CO_2 emissions, not realizing that their recommendations would still result in a net gain of atmospheric CO_2. As Sterman and Booth Sweeney (2002) explain, “The climate is complex and our ignorance great. Yet at another level it is as simple as filling a tub.”

Sterman and Booth Sweeney provide valuable extensions of heuristics and biases research into the realm of systems thinking. Erling Moxnes (2000) illustrates another systematic bias, but his research originates from a different branch of behavioral decision research called dynamic decision theory. Ward Edwards (1954, 1961, 1962) distinguishes this research, explaining that unlike static decision theory, it conceives of decision-makers as making “sequences of decisions,” which “produce both payoffs and information” (1962). In this case, a decision-maker must weigh short-run and long-run profit. He explains, “In dynamic situations, a new complication not found in the static situation arises. The environment in which the decision is set may be changing, either as a function of the sequence of decisions, or independently of them, or both.” (1962).[8]

It is this dynamic characteristic of ecological systems that gives Moxnes’ (2000) participants so much trouble. Moxnes’ study involves two different tasks—one in the context of commercial fishing and another involving the management of reindeer and lichen. In the first, Moxnes addresses the persistence of unsustainable fishing practices even when individuals were aware of overfishing. He suggests that this behavior results, not from the tragedy of the commons (Hardin 1968) as commonly believed, but from a fundamental misunderstanding of environmental systems. The participants were given a simulated virgin cod resource and "asked to build a fleet that maximized infinite horizon profits." Participants were given information regarding the previous year's harvest rates and profits. From this, they decided whether or not to add ships to their fleet. The new size of their fleet was entered into the model, and participants received the following year's harvest and profit numbers.

Despite the fact that fifty-nine of the eighty-two participants were familiar with the industry (fishermen, managers, or researchers), the median participant "built a fleet 92 percent above the fleet size that maximized the net present value given full information, and 56 percent above a boundedly rational benchmark with imperfect information" (Moxnes 2000). Only four percent of the participants were underinvested at the end of the simulation. Thus, even without pressure from competing fishermen, participants displayed a tendency to increase fishing pressure to unsustainable levels. Moxnes suggests that this behavior results from a "hillclimbing" heuristic. The individual continues to increase investment gradually until there is a decrease in catch and profits, and then decreases fishing pressure slightly. The heuristic could be successful in a static situation. However, in this case the stock of fish is being depleted, even as catch levels increase. As in the bathtub problem (Booth Sweeney and Sterman 2000), these participants fail to distinguish between the stock of fish and the flow (or catch) of fish. More importantly, reducing the stock of cod will eventually reduce the reproduction rate of the fishery. Moxnes' participants failed to account for the effects of their own actions on the fishery.

In his second study, participants exhibited a similar error from the opposite direction. They had to manage a system consisting only of reindeer and lichen. They were told that there were too many reindeer for the lichen to support and that some needed to be culled. Participants entered the number of reindeer to be killed and were then told how much lichen and how many reindeer were alive the following year. Moxnes ran this experiment three times, including once with a group of Saami reindeer herders. Even in the group of reindeer herders, which performed the best, "fewer than 50 percent of the participants were successful in improving lichen conditions within the 12-year time horizon used in the earlier experiments" (Moxnes 2000). Moreover, performances did not improve when participants were allowed to repeat the simulation, showing that they were not learning from their experience with the simulated system. Some participants looked for explanations beyond the scope of the simulation (e.g. unusually voracious reindeer, disease, etc.).

In this case, participants tended to follow a "heuristic saying that the herd size should be reduced in pace with the observed reductions in the lichen stock" (Moxnes 2000). Again, participants would make gradual changes in order to find the level of reindeer that could live sustainably on available lichen. They did not account for the fact that every year that the lichen was overgrazed resulted in a reduced lichen growth rate the following year. While the best results would be achieved by making a swift Machiavelli-like move at the outset, culling a significant portion of the reindeer and allowing the population to work back up to a dynamic balance with the lichen, participants favored much smaller steps. Moxnes suggests that the poor performance of the participants, their inability to learn, and their resulting frustration all stem from a lack of fit between the dynamics of the system and the static mental models of the participants. He concludes that "misperceptions of feedback are more devastating to human decision-making than biases in heuristics dealing with uncertainty" (2000).

John Sterman draws similar conclusions, suggesting that "[a]mbiguity arises because changes in the state of the system resulting from our own decisions are confounded with simultaneous changes in a host of other variables, both exogenous and endogenous" (1994). He also notes that delays in feedback can inhibit learning from experience. To illustrate these points to his students, Sterman (1989) developed the Beer Distribution Game. In this game four players take the roles of a retailer, a wholesaler, a distributor, and a factory. The retailer is the only player who receives direct information regarding demand. Based on this figure the retailer must order beer from the wholesaler, who orders from the distributor, who orders from the factory. There is a three-week delay from the time the player makes an order to the time it is delivered. The goal for each player is to manage their portion of the process such that money is not lost either by not being able to fill an order or by having to store an excess of beer.

Playing this game with a group of management students from MIT, Sterman set demand at four cases for the first four weeks. In week 5, he raised the demand to eight cases and held it there for the remainder of the thirty-six simulated weeks. This one change, combined with the delay in feedback, was enough to send the players into wide oscillations with their orders.[9] Sterman notes that most students blame the oscillations on "a perverse pattern of customer demand" and "are quite shocked when the actual pattern of customer demand is revealed; some voice strong disbelief" (1989). In short, suggests Sterman, "people generally adopt an event-based, open-loop view of causality, ignore feedback processes, fail to appreciate time delays between action and response and in the reporting of information, do not understand stocks and flows, and are insensitive to nonlinearities that may alter the strengths of different feedback loops as a system evolves" (1994).

In this section, we have seen that complex system dynamics adds a whole new level of uncertainty and bias that can lead to misperceptions regarding the behavior of natural and social systems. As before, two questions arise regarding the practical importance of this research.

First, are these misperceptions important? That is, do they lead to choices that result in real-world problems? And if so, then how might we correct these misperceptions and improve individuals' judgment? These questions guide the content of the concluding section.

4.5 Conclusion

Several decades ago, Jay Forrester suggested, "Our intuitive judgment is unreliable about how [complex] systems will change with time, even when we have good knowledge of the individual parts of the system" (1961, p. 14). He was writing in the context of industrial management, and he emphasized the significance of this deficiency in that context. A manager who cannot think dynamically is destined for difficulties, as illustrated in the Beer Distribution Game (Sterman 1989). Moxnes (2000) studies portend similar difficulties for natural resource managers. And certainly there are many situations in which a systems-oriented intuition of relationships and change would be an asset, just as there are many situations where knowing statistics and probabilities would result of sounder choices.

However, there is reason to believe that in light of our current environmental challenges, a basic understanding of complex systems has become particularly important. Ecologist H.T. Odum (1994) notes that humans currently put more stress on environmental systems than at any other time in history. The studies cited in the previous section—in the context of global warming and resource management—illustrate the value of a systems understanding for policymakers. However, in order for ecologically sound policies to gain public support, the public must also have some basic sense of complex systems.

Sterman and Booth Sweeney (2002) provide a good example of how a lack of systems understanding can promote complacency with regard to progressive environmental policy changes. Misperceptions concerning stocks and flows lead people to underestimate the actions necessary to decrease greenhouse gases as well as the time that must elapse after decreasing emissions before any discernable temperature change will occur. The general belief in immediate feedback and reversible processes creates a wait-and-see attitude with regard to environmental systems.

And complacency is not the only problem. A major part of the difficulty in maintaining public support for policy designed to improve the water quality of north Florida springs is the delayed feedback of the system. Water may take as long as twenty years to work its way through the aquifer and out through a spring. Therefore, the increased water quality due to changes in wastewater treatment and fertilizing practices around drainage areas may not be demonstrable for many years. The quality of water coming up from several springs is likely to worsen over the next several years simply due to previous practices. Nonetheless, one can envision how any policy implemented now could lose momentum in the face of decreasing water quality, particularly if the community was incurring costs to implement the policy. Conceivably, a public understanding of the delayed feedback would foster acceptance of ecologically sound practices. Furthermore, in that most environmental systems consist of cycles and feedback delays much longer than a three-year business projection or a four- to six-year political term, a better understanding of system dynamics could help to create a climate in which implementing policies where short-term costs were incurred for long-term benefits would not be political suicide.

But even if one agrees that a public understanding of systems concepts would create a more ecologically sustainable political climate, there is still the question of how to create a public capable of perceiving these system dynamics. Where researchers cited in the Section 3 of this chapter prescribe training in statistics and probability, systems-oriented behavioral decision researchers prescribe teaching systems concepts. Before devoting resources to such curriculum changes, it is wise to assess both the importance of changing individuals' mental models and the success of programs designed to change those mental models. With regard to importance, a critic of systems education could argue that, while Moxnes's (2000) studies clearly illustrate a deficiency with regard to systems thinking, managing reindeer is not a particularly important skill for the general public to have. After the background and examples provided in the previous two chapters, I hope that the reader understands the skills with which Moxnes was concerned and recognizes the social and individual value of such skills today. The next two chapters address the second question, regarding the success of systems-oriented education in changing mental models.

Notes

[1] Lopes (1991) refers to a *Newsweek* article (McCormick 1987) based on decision-making research. She quotes the following observation from the article as only slightly bleaker than the scientists' own conclusions: " 'Most people' are 'woefully muddled information processors who often stumble along ill-chosen shortcuts to reach bad conclusions'."

[2] Paul Meehl (1986) expresses the contextual dependence of the propriety of simplifying heuristics versus more sophisticated calculating methods, wryly suggesting, "When you check out at a supermarket, you don't eyeball the heap of purchases and say to the clerk, 'Well it looks like about $17.00 worth; what do you think?' The clerk adds it up."

[3] Kahneman and Tversky (Kahneman *et al* 1982) also identify this overconfidence.

[4] Hoch and Deighton (1989) draw similar conclusions in the context of marketing research. They suggest that when ambiguity of information is high, what one learns from experience is open to manipulation. The goal for marketers is "to interpret or resolve the ambiguity in directions that favor one's own brand."

[5] Douglas Medin and others (Medin and Bazerman 1999, Medin *et al*. 1999) suggest decision-making models that account for the social and meaning-driven components of decision-making.

[6] These models are included as Figure 4.1 and Figure 4.2.

[7] Some participants worked in problems in the context of a bank account, with deposits and withdrawals representing the inflows and outflows.

[8] For a review of dynamic decision-making experiments, see Brehmer 1992. For a review of modeling systems used in these studies, categorized by complexity, see Funke 1991.

[9] Figures 3 and 4 in the appendix illustrate the oscillations caused by the change in demand.

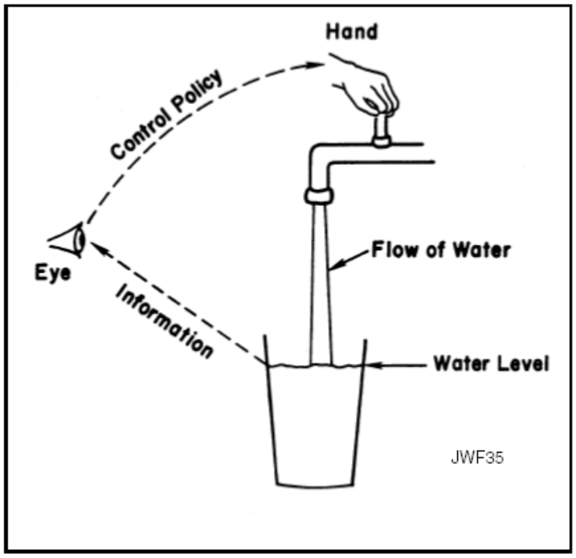


Figure 4-1. Simple system (Forrester 1996)

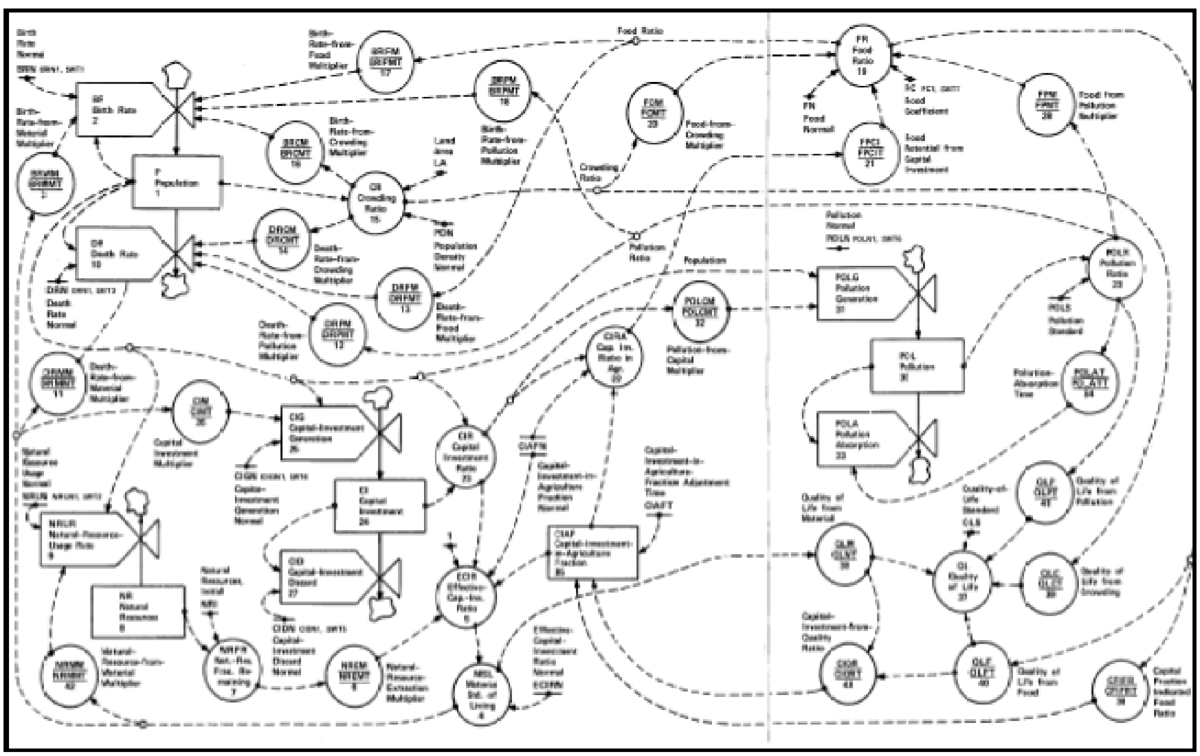


Figure 4-2. Complex system (Forrester 1996)

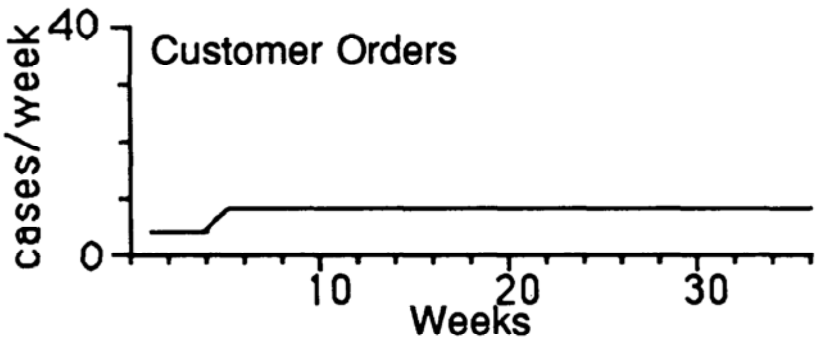


Figure 4-3. Change in demand of beer of the thirty-six simulated weeks of the Beer Distribution Game. (Sterman 1989)

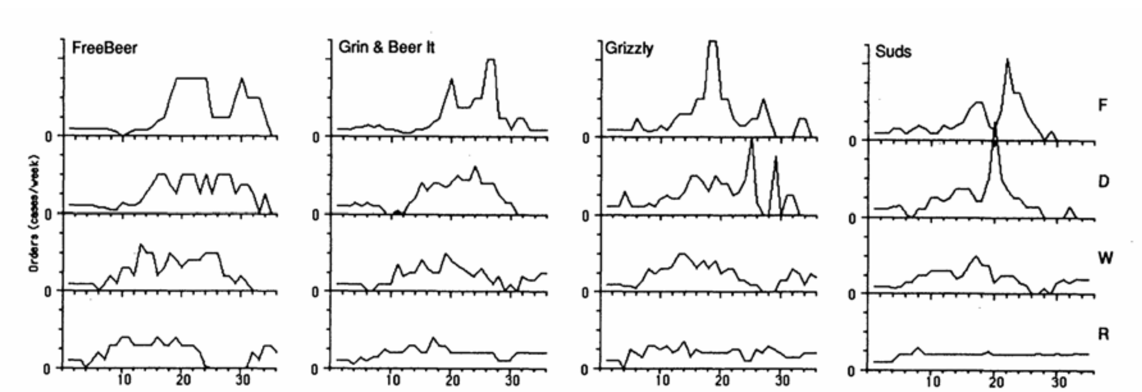


Figure 4-4. Orders recorded from four different teams in the Beer Distribution Game. Each tick mark on the y-axis represents ten cases. From top to bottom, the graphs represent orders from the retailer, the wholesaler, the distributor, and the factory. (Sterman 1989)

CHAPTER 5
COGNITIVE MAPPING AS A METHODOLOGY FOR ASSESSING THE EFFECTIVENESS OF SYSTEMS-ORIENTED INSTRUCTION

5.1 Introduction

To recap the argument so far, from Chapters 2 and 3 we might conclude that having at least a fundamental understanding of dynamically complex systems has become important both occupationally and in terms of general civic literacy. We might further conclude from Chapter 4 that whatever learning takes place during our daily experiences or from conventional curricula, it does not provide us with the skills necessary for solving even basic problems involving dynamically complex systems. Thus, there is clearly an educational need for developing these skills in our students. In this chapter we will look at the few studies that have been conducted to assess the ability of systems-oriented instruction to address this need, and from this context I will describe the methodology used in the present research.

Most assessment studies take place outside the context of formal education and assess a "systems thinking intervention"—systems workshop involving only a handful of training sessions. In one of the earliest of these studies, Huz *et al*. (1997) used systems-oriented training to help New York state health professionals from the New York State Office of Mental Health and the Office of Vocational and Educational Services for Individuals with Disabilities to integrate their operations. Instruction took place during four meetings over the course of six months, and the assessed relied primarily on participants' self-assessment of the usefulness of the systems-concepts and practices. Cavaleri and Sterman (1997) used similar techniques to assess their work with employs from an insurance company. After a handful of systems training sessions, Cavaleri and Sterman (1997) used questionnaires and interviews to measure participants' self-assessed cognitive and behavioral changes. However, this study also included an attempt to assess changes in the company's performance.

Both of these studies produced results significant in an occupational context. For example, Huz *et al*. conclude that the clients were pleased with the intervention, requesting additional interventions and that the participants themselves "perceived the intervention as productive and worthwhile" (1997, p. 163). And Cavaleri and Sterman (1997) conclude that while non-managers did not perceive significant benefits from the training, managers attribute changes in their own practices directly to the systems intervention, reporting a more "systemic" managing style. However, their attempts to assess performance changes resulting from the training were hampered by the number of confounding variables involved in the company's performance. These studies represent promising first steps toward a methodology for assessing systems-oriented instruction, but other studies were developed that relied less on participants' self-assessments.

Several researchers have assessed systems-oriented instruction using stock-flow or time-delay problems similar to those in Sterman (1989) and Booth Sweeney and Sterman (2000), discussed in the previous chapter. Kainz and Ossimitz (2002) use a pre-test/post-test design to evaluate the effect of a 90-minute course on system dynamics. Lyneis and Lyneis (2003) use a systems group versus control group design, comparing the answers of students at the end of an introductory undergraduate course in systems dynamics to those of students from a microeconomics class. Pala and Vennix (2005) collected pre-test and post-test results from both a systems group from an introductory undergraduate course and a control group consisting of students enrolled in a research design course. These studies collectively support earlier research regarding the lack of systems thinking in students who have not received systems-oriented instruction. However, the results concerning the performance of students who have received systems training is not as clear. The general trend seems to be that most systems groups show clear but often moderate improvement over pre-test or control group scores, with some systems groups actually performing worse on the post-test (Pala and Vennix 2005). The ease with which these tests can be administered allows researchers to include large numbers of participants and provide a good first look into the effectiveness of systems-oriented instruction. However, gaining insight into how systems-oriented instruction changes the way students approach problems involving complexity will require more depth than these studies can provide.

Maani and Maharaj (2004) answer this need with a study involving ten graduate students who had recently taken at least one undergraduate class in systems dynamics in a business context. Each participant was asked to work through a simulation task, acting as the CEO of a technological firm and managing five decision variables (e.g. marketing spending, price per unit) in order to meet set goals for revenue, profit, and market share. Using Verbal Protocal Analysis—which involves taping and transcribing participants who are asked to "think aloud" while working through a problem or simulation exercise—Maani and Maharaj identified several factors distinguishing the high performers from the low performers. They explain that the mere presence of systems thinking skills is not a reliable predictor of performance and suggest a more nuanced view of the role that systems thinking plays in understanding complexity.

First, Maani and Maharaj (2004) code the participants' responses using the first five skills in the typology of systems thinking skills suggested by Barry Richmond. Briefly, Richmond (2000) suggests seven different types of systems thinking with each successive one requiring all of the previous skills. The first five can be characterized as follows:

- Dynamic thinking—focus on patterns of behavior over time rather than on specific events
- System-as-Cause thinking—focus on the variables within a system that are contributing to current conditions rather than focus on external variables over which you have no control
- Forest thinking—the ability to step back from the details and look at the entire system

- Operational thinking—getting away from the lists of “key drivers” that contribute to a system’s behavior and focusing more on the structure of relationships to see how the “key drivers” affect the system’s behavior
- Closed-loop thinking—looking for the connections that turn linear causal chains into causal loops[1]

Maani and Maharaj report that while the presence of dynamic thinking and system-as-cause thinking did not necessarily predict high performance, the presence of higher order thinking skills—specifically, forest thinking, closed-loop thinking, and operational thinking—did seem to play a major role in participant performance.[2] In light of these results, future studies might involve looking specifically for these higher-order skills. However, since Verbal Protocol Analysis is highly labor intensive, participant numbers must be kept relatively low (between 5 and 20). A more broadly applicable methodology would allow one to identify these skills more quickly so that more participants could be included in a study. Cognitive mapping methods may represent a useful middle ground between the studies discussed in this section by providing deeper insights into participants’ thinking than the stock-flow and time-delay questions of the studies above while including more participants than Verbal Protocol Analysis allows.

5.2 Cognitive Mapping

In 1948, psychologist Edward Tolman described two competing theories regarding learning in the context of laboratory rats navigating through a maze. According to the first theory, stimulus-response, the rat’s brain is analogous to a room of telephone operators, receiving incoming calls from sensory organs and sending outgoing messages to muscles. Learning in this view is a matter of strengthening the connections to appropriate responses. According to the second theory, which Tolman called field theory, learning involves the establishment of “something like a field map” (Tolman 1948, p. 192). In the second theory—the one to which Tolman subscribes—the rat’s brain is viewed as a map control room, in which stimuli, if allowed into the room, "are worked over and elaborated… into a tentative, cognitive map of the environment (p. 192). Tolman describes a series of experiments, which, he suggests, support the field theory characterization of the rats' behavior.

This study by Tolman not only marks the beginning of the broad field of cognitive mapping, but also introduces a key distinction in the context of the current study. Tolman suggests the importance of discovering the breadth of the rats' cognitive maps, splitting the maps into two types: strip maps and comprehensive maps. A strip-map refers to a linear set of connections, which in the context of Tolman's experiments, might look like a set of directions—turn right, go forward, turn left, etc. These directions, if correct, can lead a rat successfully to the food at the end of the maze. A comprehensive-map, as the name suggests, reflects a broader understanding of the maze. Rats with more comprehensive maps, suggested Tolman, were better able to respond to slight changes in the maze. Tolman even reports of rats "who, after having learned an alley maze, pushed back the cover near the starting box, climbed out and ran directly across the top to the goal-box where they climbed down again and ate" (p. 203). The significance of strip-maps versus comprehensive maps will be clear shortly. For now, let us take a look at more contemporary cognitive mapping techniques.

5.3 Contemporary Cognitive Mapping

Research with cognitive mapping can be split into two main categories. The first category, commonly represented in the fields of psychology and neuroscience, includes studies designed in the same vein of Tolman's original studies with the focus on cognitive representation of physical environments. In recent years, researchers have investigated the spatial orientation of rats (Cohen and Bussey 2003), gerbils (Ellard 2000), pigeons (Bingman et al. 2006), goldfish (Jacobs 2003), honey bees (Fry and Rudiger 2002), European badgers (Bodin and Benhamou 2006), and Guinean baboons (Llinas et al. 2003), just to name a few. Explorations into how humans develop spatial orientation also represent a rich branch of study within this first category (e.g. Moeser 1988; Portugali 1996; Pentland et al. 2003).

The second category differs from the first in that the mental representations under study no longer correspond to a physical landscape. In these studies, the maps can be used to explore participants' representations of processes or situations—what we might call conceptual landscapes—ranging anywhere from caring for disabled family members (Shewchuk et al. 2004) to making managerial decisions regarding new products (Tyler and Gnyawali 2002). Instead of orienting people in their geographical environments, maps, in this broader context, orient people in their informational environments. In this context, Anne Kearny and Stephen Kaplan (1997) define cognitive maps as "hypothesized knowledge structures embodying people's assumptions, beliefs, 'facts,' and misconceptions about the world" (1997, p. 580). The shift from physical to informational landscapes has enabled much wider applicability of cognitive mapping tools across a number of different fields, including political science (Axelrod 1976), management studies (Eden 1992), and computer science (Kardaras and Karakostas 1999).

Not surprisingly a wide array of cognitive mapping techniques has been developed in the context of this second category. I do not attempt to cover them all here. My aims are first to establish the potential for cognitive maps as effective tools for studies regarding environmental issues and education, and then to provide a brief explanation of some existing cognitive mapping techniques and concepts. Finally, I will explain how these techniques have been adapted to assess the value of systems-oriented instruction.

5.4 Cognitive Mapping as a Tool for Environmental Education

As the field of environmental education evolves, researchers are forced to ask more and more complex questions in an attempt to understand the indirect causal relationships between educational programs and students' perception of the environment. Focus has shifted from what information students have acquired about the environment to how students understand and feel about environmental systems and their places within those systems. Questions of perception become difficult to manage when dealing with a wide array of stakeholders involved in a complex environmental issue. While some researchers (Kearney et al. 1999) have suggested that cognitive mapping can serve as an effective tool for gaining a sense of people's perspective concerning environmental issues, little of this type of research exists.

Many of the questions guiding cognitive mapping studies in other fields parallel those regarding environmental issues and education. For example, Clarke and Mackaness use a form of cognitive mapping to study decision-making by upper-level management, noting that these managers often find themselves "in uncertain situations where they are faced with insufficient facts and complex alternatives" (2001, p. 148). Upper-level managers are not unique in this regard. Any individual attempting to understand contemporary environmental challenges faces the same circumstances. Clark and Mackaness (2001) explain that high-level managers often rely on their "gut feeling," making intuitive decisions, and they suggest that this "intuition" can be identified based on a manager's cognitive map. Presumably, the same would be true in the context of environmental decision-making.

In another study, Jenkins and Johnson (1997) suggest that the most striking difference between high-performing and low-performing managers (as measured by profit) is the perceived connections between three categories of concepts: actions, customers, and performance measures in their maps. While low-performing managers included these concepts on their maps, they did not connect them, as did the high-performing managers. Jenkins and Johnson use this observation to suggest that such connections are important for effective management.

The application of this type of inquiry regarding environmental issues—an area of study based largely on connections between individuals and their environment—should be clear. For example the ability to include categories like rivers, rain, wetlands, and faucets on a map about water may not indicate any understanding of one's dependency on water resources. The ability to express connections between these categories may be a stronger indication of one's understanding of—and perceived attachment to—those resources.

Following the implicit analogy to geographical maps may help to clarify the potential for cognitive mapping as an environmental research tool. Echoing Tolman's reference to strip maps and comprehensive maps, Fiol and Huff (1992) discuss the distinction between using a roadmap or written directions to find a particular location. First-time visitors to a place may be able to recount the directions they followed to get there, but are less likely to be able to draw a map or the surrounding area. If the goal of a study is to assess participants' ability to get from point A to point B, evaluating the written directions would be enough. However, if a researcher wants to learn more generally about participants' understanding of the area, the written directions would not be the appropriate tool.

The information contained in the written directions is analogous to what one could acquire from surveys or questionnaires. Indeed when asking relatively concrete research questions (e.g., Can the participants get from point A to point B? or Do residents know at least five precautions to take to protect their homes from wildfires?), then these types of tools are probably the most effective. However, when the questions are somewhat more abstract and focused participants' deeper understanding of the situation in question (e.g., How might we characterize the participants' understanding of the terrain between point A and point B? How might we characterize residents' understanding of wildfire?) cognitive mapping proves to be an effective tool. Kearney and Kaplan make a similar point, emphasizing the concept of ownership and calling for researchers to distinguish "between a 'purely verbal understanding' [of a concept] and a deeper grasp of the concept and what it refers to" (1997).

Additionally, cognitive mapping seems particularly appropriate for looking at interpretations of complex social and ecological systems because experts use similar types of maps to describe and/or model these systems. The structure of maps used in both social network analysis (Johnson et al. 2001) and ecological modeling (Fath and Patten 1999; Ulanowicz 2004) parallel the structure of cognitive maps. Thus, a researcher can evaluate cognitive maps with tools and indices similar to those used to evaluate the systems themselves.

But claiming that cognitive mapping could be an effective evaluation tool is a little like saying that a wheeled vehicle could be an effective mode of transportation. In truth there exists a multitude of approaches to cognitive mapping, each adapted to the researchers' particular needs, limitations, and interests. Huff (1990) explains that one reason for the diversity of approaches is the diversity of relationships between two cognitive elements A and B:

> Simple connotative association (A reminds me of B); degree of similarity (A and B are different); relative value (A is more important that B); and causal linkage (A causes B) might be identified…. Or, one might map arguments (since A is true, B is not true); focus on choices (since A, we must do B); and make inferences beyond the relationships present in the text (since A and B are mentioned, the informant must be influenced by C). (Huff 1990)

If one considers cognitive mapping to be the genus, then the following three sections provide a description of three species of cognitive mapping—concept maps, cause maps, and conceptual content cognitive maps—that represent much of the work currently being done.

5.5 Concept Maps

Novak defines a concept as "a perceived regularity in events or objects, or records of events or objects, designated by a label" (1998). If this definition sounds somewhat convoluted, it is because Novak tries to account for the fact there are many concepts that, while difficult to define, are easy to describe. The major premise behind concept maps (although this premise could support other forms of cognitive mapping as well) is that our familiarity with a concept can be measured largely by how many propositions we are able to attach to that concept (Novak and Gowin 1984). For example, while we may be uncomfortable defining “water” in any scientific or unique way, we are familiar enough with the concept to express number of propositions concerning water (e.g., Living things need water; water freezes at 0°C; water collects as clouds in the sky). The purpose of a concept map, then, is “to represent meaningful relationships between concepts in the form of propositions” (Novak and Gowin 1984). Figure 5.1 shows an example of a concept map of water provided by Novak.

One interesting aspect of these maps is that the hierarchy can shift depending on one’s perspective. For example the map in Figure 5.1 on water can become a map on living things with water, plants, and animals as the subordinate categories. Also, note the “determines” arrow from the “motion” node to the “states” node. This is significant in that not only can the map show hierarchy, but it also has the ability—although somewhat limited—to show lateral relationships.

Evaluation of concept maps consists of looking at the number of nodes, length of chains, and number of different hierarchical levels. However, Novak (1998) discusses the evaluation of these maps like one might discuss grading an essay. Rather than specific indices to calculate, one would look for hierarchical differentiation just as one would look for proper organization of ideas in a piece of writing. Nonetheless, many of the indices mentioned above could apply.

The difficulty involved with using concept maps as a research tool is that, as Novak (1984; 1998) presents them, they are a skill that students acquire over time. The context here is that of a teacher asking students, who already have the skill of producing their own maps, to create a map about a curriculum unit that they have just covered. The complexity of the map, argues Novak, would be a better indication of the student's understanding of the new concept than would a typical classroom test.

The added time necessary for students to learn the new skill may not present a problem for researchers or teachers evaluating their own long-term program. However, one can imagine how the mapping skills required on the part of the respondents could cause problems for evaluators who have only limited time with the respondents. Without devoting sufficient time and effort into learning the skill, a respondent's failure to produce a complex map could be the result of unfamiliarity with concept mapping rather than with the concept in question. Still, one could try to incorporate some of the more useful aspects of these maps into other mapping techniques.

5.6 Cause Maps

Cause (or Causal) maps (an example of which is shown in Figure 5.2) represent another popular form of cognitive mapping. As the name suggests, the maps represent the participant's perception of causal relationships between items on the map. Most methods follow the example presented by Axelrod, in which the researcher transcribes a meeting or an interview and codes the text "sentence by sentence or even phrase by phrase" (Axelrod 1976). Axelrod concedes that such a procedure "requires a large number of subtle coding decisions," but explains that "after more than three years of work, the coding rules [which comprise a forty-page appendix to the text] have reached a state of precision such that the intercoder reliability is fully compatible with the accepted standards of good quantitative work in the social sciences" (1976). Some researchers (e.g., Clark and Mackaness 2001) include an extra step during which the respondents validate the maps.

Other researchers have the participants make the maps themselves. Jeffrey Ford and W. Harvey Hegarty (1984) explore management beliefs of MBA students by having the students make a causal map using ten concepts provided by the researcher. While this technique for eliciting the maps avoids the coding step, dictating the concepts included on the map takes away from the participants' ability to show their unique picture of the situation in question. Participants may be using concepts that they feel are unimportant or, conversely, the researcher's chosen concepts may omit something that is particularly important to the participant.

One technique for avoiding coding while still allowing participants to choose their own concepts is the Self-Q Interview (Bougon *et al* 1983). This method actually requires three interviews for each respondent. In the first interview, the interviewer asks the participant to make up her own list of questions regarding the topic of interest. In Bougon's example, teachers are asked to create a list of questions regarding the classroom experience. Before the second interview, the interviewer creates "notion cards" based on each participant's questions. These notion cards will later become nodes on the maps. Creating notion cards basically consists of converting questions to assertions. For example, "Will the students come to class on time?" becomes "The students come to class on time."

In the second and third interviews, the respondents review their questions and the notion cards to make sure that the notion cards express the ideas intended by their questions. The respondents then choose the eleven most important nodes[3] are guided through a list of steps during which they identify existing relationships between those eleven nodes, labeling those relationships as positive or negative. Note that in this context "positive" and "negative" does not signify likes or dislikes. For example, the statement, "The more I give detention for tardiness, the less students arrive late for class" expresses a negative relationship. The first notion—giving detention—has the effect of decreasing the second notion—students arriving late. The fact that the second notion is desirable is immaterial in this context. With this information, maps can be constructed by connecting the nodes with arrows to signify causal connections and labeling the arrows positive or negative as the respondent indicated. The primary benefit of this technique over Axelrod's is the participation on the part of the respondent in developing the map itself. The interviewer has very little influence on what the respondent's map will look like.

Jenkins and Johnson (1997) combine both Axelrod's and Bougon's techniques, creating a set of cards (similar to notion cards) from the first interview and conducting second interviews to discuss the items on the cards. Transcripts from the second set of interviews are coded following Axelrod's (1976) example using the items on the cards as the nodes of the map. This method marks an improvement over Axelrod's in that the transcribed interviews are more likely to be focused on the concepts in question, facilitating coding. However, the intensiveness of this method precludes large sample sizes.

Evaluation of causal maps depends largely on what sort of questions one is asking. Eden *et al* (1992) suggest an assortment of indices that may provide some measure of the complexity of an participant's map. First, one might simply tally the number of nodes or constructs included on the map, with the assumption that more constructs may suggest a more developed understanding of the issue (or at least the participant's belief in a more developed understanding). Second, one could calculate the ratio of links to nodes. The idea here is that a high ratio would indicate more interconnectedness and, consequently, "a greater understanding of the relationships impacting an issue" (Jenkins and Johnson 1997). The average length of causal chains may also be indicative of the informant's understanding of the direct and indirect effects inherent in an issue.

Eden *et al* (1992) also suggest a domain analysis, which acts as a measure for the relative importance of a particular node to the participant. In this analysis, one calculates the total number of in-arrows and out-arrows from each node. One can also broaden this analysis by adding successive layers of domain. Clustering software offers another level of analysis that could help one to identify the "most robust parts of a cause map—those parts of the map that are relatively insensitive to small changes in the structure of the map" (Eden *et al* 1992).

However, these indices are not always informative. In Jenkins' and Johnson's (1997) study, these indices did not show any significant difference between low-performing and high-performing managers. The researchers were able to identify a consistent difference only by looking closely at the content of the connections being made. Thus, the presence or absence of specific connections may also be an important part of an evaluation.

5.7 Conceptual Content Cognitive Maps

Anne Kearney and Stephen Kaplan (1997) present a slightly different mapping technique, called Conceptual Content Cognitive Mapping (3CM), designed specifically "for measuring people's perspectives on, or cognitive maps of, complex domains" (1997, p, 599). Unlike concept mapping, 3CM does not require training on the part of the participants. With this technique, participants are introduced to the topic or issue of study and then asked how they would explain that issue to a friend who knew nothing about it. More specifically, they are asked to list the "things" or "aspects" of the issue that they think are important. The concepts listed by the participant become the nodes of the map. The participants then group the nodes and label each of the groupings with a short, descriptive phrase.

As shown by the example in Figure 5.3, the 3CM technique—unlike causal maps—does not establish causal relationships between nodes. Instead, the technique produces a hierarchical structure of items and categories that provide insight into the associative links perceived by the participant. In a sense, 3CM is not entirely unlike an abbreviated form of the Self-Q test. By having respondents either develop their own nodes or choose them from a stack of cards, the researcher ensures that only those aspects which the respondent deems relevant will be included on the map. As Kearney and Kaplan (1997) explain, "[W]e would expect participants to be able to differentiate objects they own from those they do not."

The obvious advantage of 3CM over the Self-Q Interview is the relative ease of the technique both for the interviewer and the respondent. Also, the categories supplied by the respondent during 3CM provide the researcher with a picture of how the respondent would aggregate the nodes into a hierarchy. And lastly, because of the relative ease of developing and handling these maps, respondents are encouraged to provide as many nodes as possible. Consequently, a map can contain well over ten nodes and still be manageable. However, the disadvantage to the 3CM technique is the lack of lateral connections between modes. The technique tells the researcher which nodes are perceived to be closely related, but does not provide any information regarding the quality of those relationships.

These maps lend themselves to many of the same evaluation indexes already discussed. In addition, these maps lend themselves to the formation of meta-categories or themes. For example, in a study on public perception of wildfires, one can calculate a measure of the relative importance of fire prevention by calculating the proportion of nodes that could be included within this theme.

5.8 Interview Methods

This study combines 3CM methodology with that of causal maps. Using 3CM techniques for soliciting concepts ensured that participants would be working only with those concepts they felt were important, while the inclusion of the causal mapping step provided an illustration of how the system in question was organized in the eyes of the participants. By combining these techniques, I am able to ascertain which aspects of the system are most important to a participant and what associative and causal connections the participant places between those aspects.

5.8.1 Developing a Scenario

3CM mapping is generally used in the context of a subject with which participants are already familiar. For example, Kearney et al. (1999) used 3CM techniques to study stakeholder perceptions of forest management. The participants, who included environmentalists and members from the USDA Forest Service and the timber industry, were all deeply familiar with the issue at hand. However, for my own study the mapping exercise begins with participants reading an article about a topic with which they are unfamiliar. This modification represents a significant shift of focus made necessary by the goals of the study.[4] In their forest management study, which is representative of most studies using 3CM, Kearney et al. (1999) explore participants' perceptions regarding a specific topic. In contrast, the goal of the present study is to explore a critical thinking skill (i.e., systems thinking). Certainly a specific topic would still be necessary in order to provide participants with a context in which they might use their thinking skills. However, the topic itself was a means rather than an end.

Nonetheless, the topic of the article needed to be carefully chosen. In order to be useful for this study, the topic needed to meet three requirements. First, it needed to include complex relationships, involving natural and social systems. The motivation for this study stems from that fact that this characteristic—common to environmental issues—is the reason for much of our difficulty in understanding and addressing our most pressing environmental challenges. Second, the topic needed to be unfamiliar to the participants. In an attempt to follow the convention of choosing topics familiar to participants, I designed early pilot tests of this research using global warming as the subject for cognitive maps. While some interesting differences were identified in these maps, most of these differences failed to provide insights regarding thinking skills. People understood that the burning of fossil fuels increases atmospheric levels of greenhouse gases, thereby increasing global temperatures. The inclusion by some of other relevant processes (e.g. volcanic gas emissions, loss of forests) produced some variability, but these differences seemed to indicate a difference in knowledge level rather than a difference in the thinking process itself. Thus, while the data might have been useful in a study specifically about people's perceptions of global warming, they did not appear promising for this study.

The one difference of interest in the context of this study was the inclusion by some of feedback loops on their causal maps. That is, while most participants depicted a series of causal connections that included an increase in fossil fuel burning leading to an increase in greenhouse gases, which then led to any number of undesirable events (e.g. desertification of farmland, severe weather, sea level rise), some participants also included policy or education concepts that resulted in a decrease in the burning of fossil fuels. The significance of the inclusion of feedback loops will be discussed further in the following chapters. For now, suffice it to say that the inclusion of such connections can be interpreted as an awareness of actions that might decrease global warming and/or a belief that such a decrease is possible. Thus, while the global warming pilot tests provided some promise that the tool would be able to measure differences attributed to systems thinking, it became clear that in order to explore a thinking skill, an unfamiliar topic must be chosen for the exercise.

The third requirement of the topic, paradoxically, was that it needed to be something to which the participants did have some connection. That is, the topic must be both comprehensible and consequential. Using, for example, the topic of complex brain processes or water management practices in rural Thailand—both of which meet the first two requirements—would likely discourage active participation in the exercise, either because they are too difficult or too foreign to inspire active participation.

With these three requirements in mind, I elected to use the controversy over the Atlantic menhaden fishery as the topic for this study. Menhaden, a relative of the herring, is a bony, oily fish with high economic and ecological value. The controversy involves a web of ecological, economic, and social interactions that make it ideal for exploring how people interpret complex systems. Also, despite its importance, few people are familiar with the controversy. As an extra precaution, I changed the name of the fish in the article from menhaden to samaki—a fictitious name—and made a few other minor modifications, ensure that prior knowledge about the specific issue would not be a factor.[5] And finally, although most people are unaware of the controversy over menhaden, most participants can easily relate and identify with catching and consuming fish and fish products. In short, the scenario involves a high level of complexity and is obscure enough to ensure that no one had previous knowledge of the system, while being familiar enough to allow people to identify with the subject matter.

The tone of the two-page article describing the samaki controversy is intended to mimic that of a newspaper. However, the density of information in the article is decidedly higher. The intention here is to overwhelm the participant—at least mildly. Since total recall of the information in the article is unlikely, participant maps will likely include only those concepts that the participants consider familiar and/or important.

5.8.2 Creating Concept Cards

There are two forms of 3CM techniques—open and structured. During the open technique, participants are prompted to produce a list of concepts pertaining to the topic at hand. The researcher records these concepts as the participant says them. During the structured technique, the participant chooses concepts from a list of cards provided by the researcher and is then given the option to add more cards if necessary. The benefit of the open form of 3CM is that it ensures that only concepts that come directly from the participant are included in the participant's map. However, because of the idiosyncratic nature of the responses, this method does not lend itself to large sample sizes and introduces a degree of subjectivity when evaluating and comparing participants' maps. In early pilot testing for this study, the subjectivity was exacerbated during the causal mapping portion of the exercise.

For these reasons, I used the structured form of 3CM. The concept cards were developed over many stages of pilot testing. Early pilot tests involved the open form of 3CM, during which most of the concepts for the cards were identified. Concepts were also added during subsequent pilot tests using the structured 3CM. If when given the opportunity, pilot test participants identified concepts that were not included in the concept cards provided, that concept would be added to the list of concepts for subsequent tests. I made no judgment about the relevance of a concept that a pilot test participant wanted to add. If the concept was not, in my opinion, already represented in the cards, then it was added to the list. By the time the pilot testing was over, thirty-six concepts had been identified. This fits into the range of 30-50 suggested for 3CM by Kearney and Kaplan (1997, p. 605).

The structured method does present some potential problems as well. By providing participants with a list of concepts, I am dictating more specifically how participants will be asked to think about these concepts, and I am providing the opportunity for participants to include concepts that they would not have come up with on their own. I address this problem in a number of ways. First, the high density of information in the article described earlier will encourage participants to choose only those concepts that appear most significant. Second, the inclusion of the causal maps decreases the likelihood of a participant choosing an unfamiliar concept because they are forced to identify that concept's role in the larger system. If a participant was unable to causally connect a concept to any other card, then he was asked to eliminate the concept from both the 3CM and the causal map. In addition, as Kearney and Kaplan (1997) point out, because the list is generated by earlier participants (rather than by experts) it is more likely to include aspects of the system that are salient to the participant. In other words, participants using the structured 3CM method are not forced to think like experts.

5.8.3 3CM Exercise

The cognitive mapping sessions were conducted in groups of 20-30 students with each participant working individually. The exercise lasted between 60 and 80 minutes. Participants were given ten minutes to read the article. They were allowed to take notes and to reread the article if time allowed. Participants were allowed to ask for clarification regarding the definition of a word, but were not allowed to receive any other help with the article. At the end of the ten minutes, the article was taken away. Participants were then asked to go through a stack of thirty-six cards, with each card expressing a concept pertaining to the article. After the participants identified the concepts they thought were important and laid the rest aside, they were given the opportunity to write in any relevant concept not listed in the cards.

Once the participants had included all the concepts deemed important, they were asked to sort the concepts into groups based on whatever criteria they chose. No restrictions were put either on the number of cards chosen or on the number of cards in a group. Participants were then asked to label the groups with short descriptive titles that indicated why those cards were together. After labeling their groups and reviewing their groupings, participants moved on to the causal mapping portion of the exercise.

5.8.4 Causal Mapping

For the causal mapping portion of the exercise, participants were given instructions about how to show causal relationships between their cards. In short, the participants were asked to identify the relationship between any two cards as fitting into one of the following categories:

- An increase in Card 1 leads to an increase in Card 2.
- An increase in Card 1 leads to a decrease in Card 2.
- An increase in Card 1 does not affect Card 2.

Pairs described by the first sentence were connected with positive or "S" arrows, designating that they move in the same direction. Pairs describe by the second sentence were connected with negative or "O" arrows, designating that they move in the opposite direction. And pairs described by the third sentence were not connected by arrows. Note that the arrows were directional. Participants were encouraged to assess the same pair of cards, transposing Card 1 and Card 2.

A simple, four-card example was used to model the use of positive (S) arrows, negative (O) arrows, bi-directional causality, and feedback loops. The use of an example during this part of the directions was intended to model the inclusion of non-linear connections and feedback. Some causal mapping studies (e.g. Roberts 1976; Eden and Ackermann, 1998) use a pairwise method for comparison, rather than having participants generate their own maps. With this method, participants are asked to attach sentences similar to the three listed above to every possible pairing of the cards they have chosen. I deviated from those methods for two reasons. First, from a purely logistical standpoint, the number of cards included in this exercise made explicitly addressing each possible pairing of cards impossible. The average number of nodes chosen for each group ranged from 11 to 20, giving the average number of possible pairings a range of 65 to 210. Some participants included over thirty cards, meaning they would have had to consider several hundred pairings. Bougon (1983) addresses this problem by limiting the number of cards, but this solution also results in at least some of the participants not being able to produce a complete map.

More importantly, forcing participants to consider each relationship explicitly would have defeated the purpose of the assessment. The hypothesis for the current research, stated generally, is that students untrained in systems thinking will be less likely to look for non-linear causal links and feedback loops. The question, in other words, is not whether students, when forced to consider a connection, can recognize it. Rather, it is whether students will look for—and identify—such connections on their own. In a study comparing these two methods for eliciting cognitive maps, Hodgkinson et al. report that "the pairwise comparison method included just greater than 5 times more causal links between nodes than maps derived from the freehand method" (2004, p. 17). Clearly, the pairwise method encourages participants to include connections they would not otherwise consider. Therefore, all the tools the students need are modeled in the small example, and the students are encouraged to look for all possible connections, but they are not forced to do so. Doyle et al. (1998) warn that when attempting to measure systems thinking, the assessment tool itself can cause participants to approach the topic in ways they would not have considered otherwise. The example was designed to minimize this danger while ensuring that each participant knew that non-linear causality was allowable in the context of the exercise.

After the example was presented, the participants were asked to follow the same conventions to complete their own maps about the samaki controversy. During this part of the exercise, participants were allowed to ask for clarification regarding the meaning of an arrow. For example, a participant could be told that what they had drawn implied that an increase in the samaki population leads to a decrease in the sport fish population, but a participant could not get any feedback regarding the accuracy of the connection. Once participants had completed their maps, they were asked to review it once more, checking to see that the map accurately represented their understanding of the situation. This completed the mapping exercise.

5.9 Conclusion

The next chapter includes discussion about how one converts the maps produced during this exercise into quantitative data. But before moving on to the evaluation of the data, it is worth pausing for just a moment to consider the pedagogical value of the exercise just described. While it was designed primarily as an assessment tool, the exercise has proven to be valuable as a tool for encouraging critical thinking about the situation in question. Novak (1988) writes extensively about the use of concept maps in the classroom, but these techniques have limited popularity among educational practitioners. The inclusion of causal maps appears to be even less common, which is unfortunate in light of comments made by many of the study participants. The visual and tactile aspects of moving the cards around and making a picture of the situation seems to be inherently appealing. It was not uncommon for participants to express their enjoyment of the exercise as the fuzzy image in their minds became more concrete on the map in front of them. These comments are consistent with Ambrosini’s and Bowman’s description of causal mapping as “a simple but yet powerful technique that can help us in surfacing tacit skills” (Ambrosini and Bowman 2001, p. 817) and with Roos’s and Hall’s (1980) suggestion that causal maps can help individuals identify feedback loops within a system.

Doyle et al. (1998) suggest that assessment and education are mutually exclusive goals when determining the affect systems thinking has on students’ interpretation of events. Their concerns are sound. If a tool intended to measure systems thinking actually teaches systems thinking, then the accuracy of the measurement may be compromised. Maintaining a distinction between assessment and education requires, first, a method of eliciting the maps that does not require extensive training. The exercise must be designed such that anyone can follow the directions provided. Second, any discussion of the system being mapped must be postponed until after everyone has completed the exercise. At this point the tool can shift from an assessment tool to an educational tool. As such, the maps can be used as a basis for class discussion about the topic in question or as a pre-writing step that helps students to develop a clearer picture of their own thoughts before trying to organize them into a piece of writing. And finally, the maps can be used to teach the same skills that they were initially used to assess. Themes within the maps and patterns of non-linear dynamics can be identified and evaluated in order to foster critical thinking in general and systems thinking in particular.

It is perhaps fitting to close this chapter by returning to Tolman and his rats. In the same 1948 study that sparked the field of cognitive mapping, Tolman relates an experiment particularly significant in the context of the present research. In this experiment, a rat was placed in a cage with a specific striped visual pattern above a food dish wired to have an electric charge. A rat that attempted to eat from the food dish was zapped. Consequently, the rats quickly learned to associate the visual pattern with the electric shock. Tolman reports, "when the rat was replaced in this same cage days or even weeks afterwards, he usually demonstrated immediately strong avoidance reactions to the visual pattern" (p. 200). Common among these avoidance behaviors was piling up sawdust in order to cover the visual pattern. Such a response is not entirely dissimilar from Sterman and Booth Sweeney's (2002) or Moxnes (2000) participants discussed in Chapter 4, who, lacking an understanding of the system's, addressed challenges superficially, irrationally, and with no more success than the rats piling sawdust.

Notes

[1] Richmond's last to systems thinking skills are quantitative thinking, which involves putting rough quantitative values on things that are impossible to measure (e.g. motivation, self-esteem), and scientific thinking, which refers to a general willingness to change or discard one's model of a system as new evidence dictates rather than stubbornly trying to defend one's own view as "truth."

[2] Maani and Maharaj (2004) also discuss the importance of each participant's approach to understanding the system at hand. Even the use of higher order systems skills are not enough to ensure high performance. The highest performers used these skills to learn about the system through iterative cycles of conceptualizing, planning, and acting.

[3] The number of cards is restricted to a manageable number in the context of the remaining steps of the exercise.

[4] While the use of an article is a departure from other studies using 3CM, it is not unheard of in the field of cognitive mapping. For example, in an exploration into how managers form an investment strategy Hodgkinson et al. (2004) provide participants with a "short case vignette…concerning a strategic investment decision" and a brief summary of the company's history.

[5]Arguably, this precaution was unnecessary, as several of the pilot testers who went through the mapping exercise before the change to samaki had thought I made up the name menhaden.

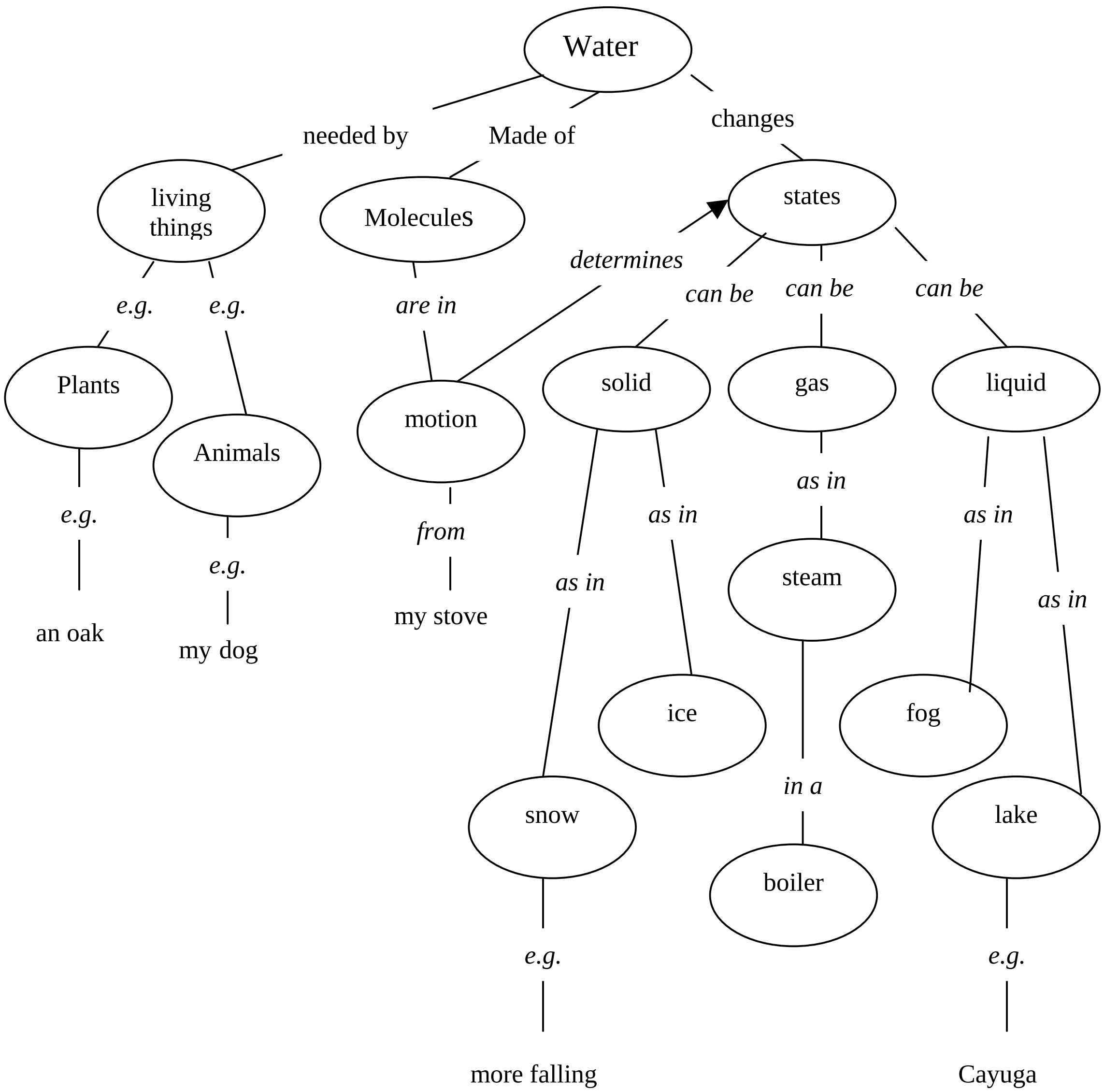


Figure 5-1. A concept map for water taken from Novak (1984).

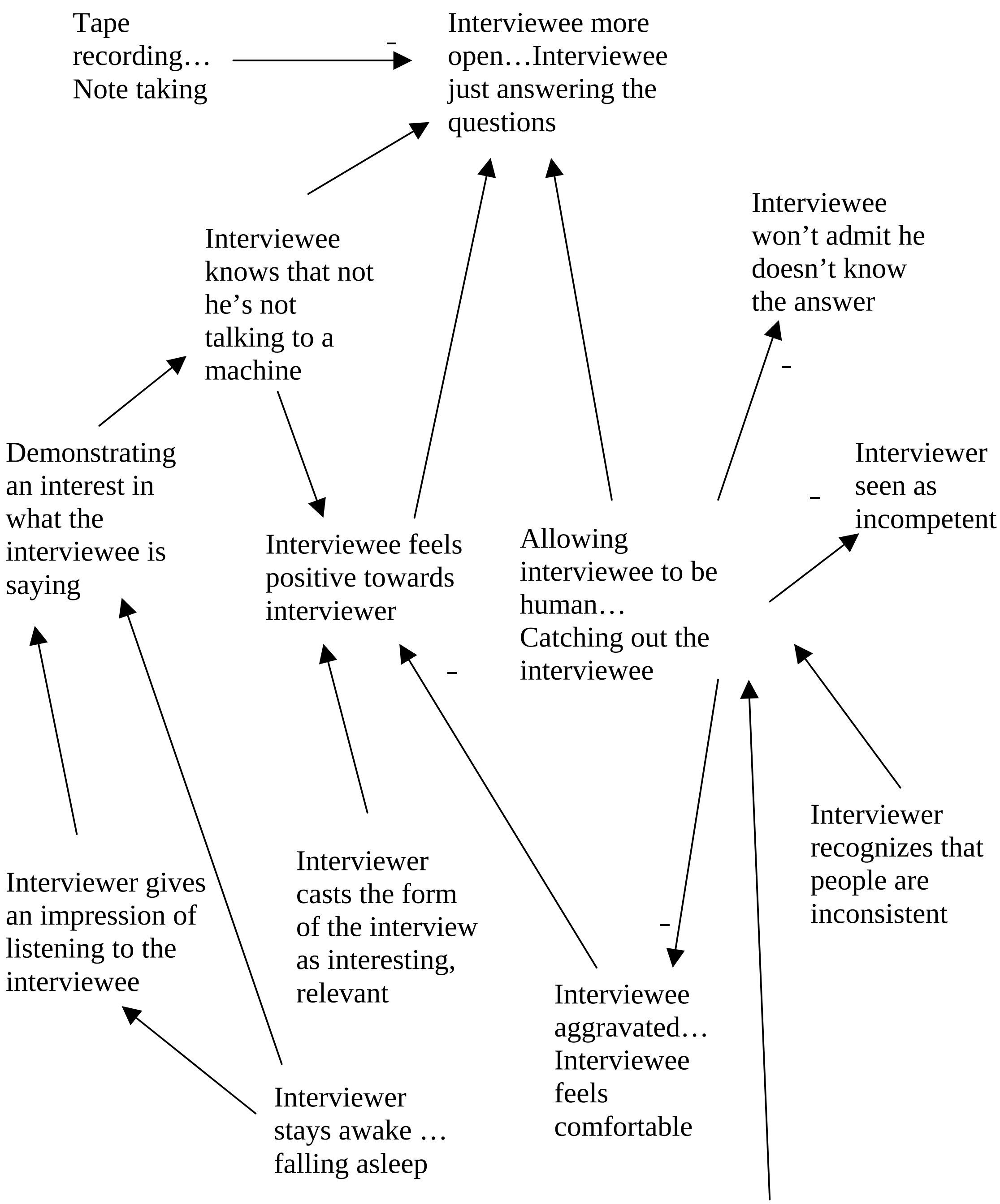


Figure 5-2. A cause map depicting interactions in an interview (Eden et al, 1992). Note that "…" should be read as "rather than."

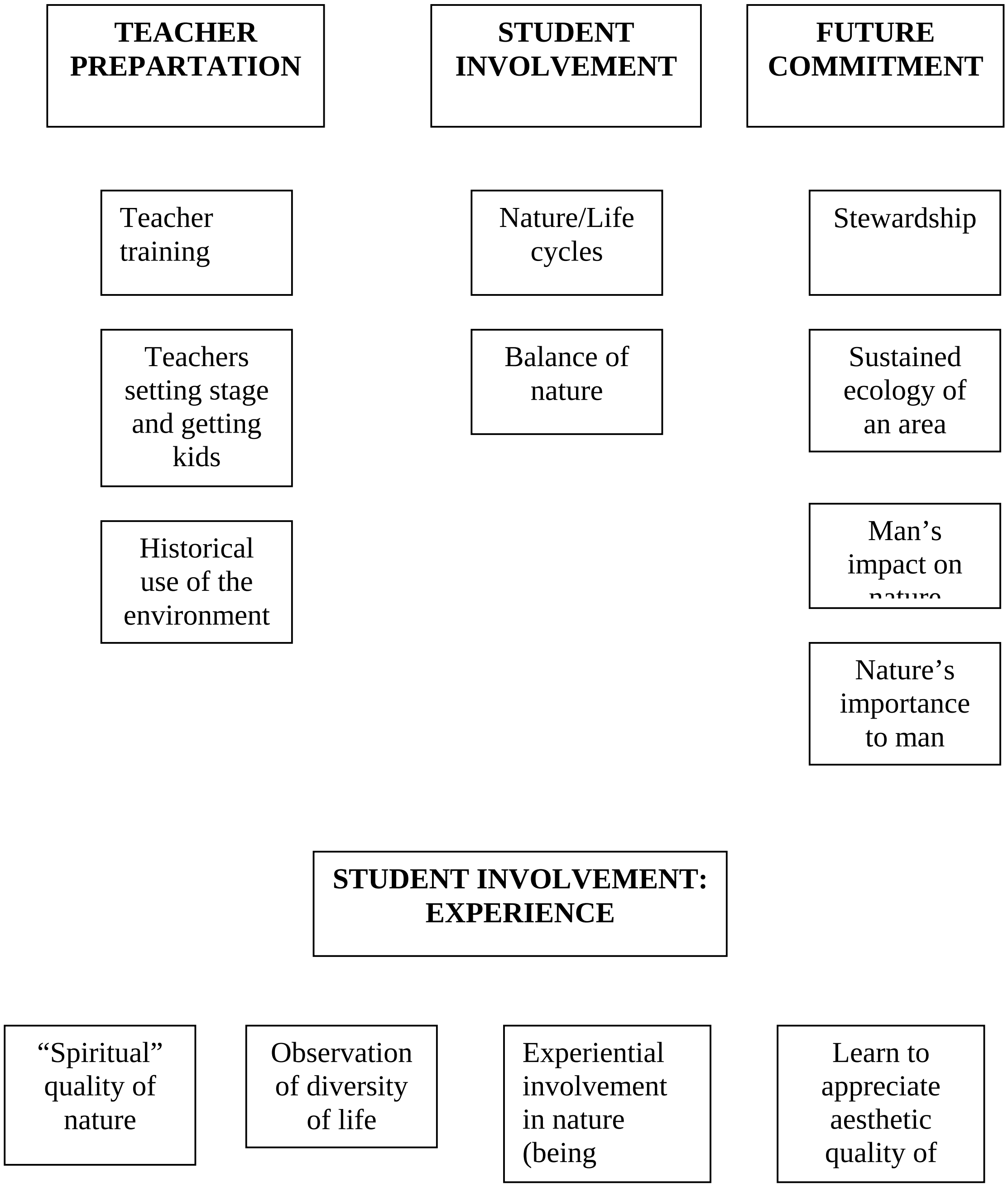


Figure 5-3. A 3CM map representing a non-EE teacher's view of environmental education (Kearney 1999)

CHAPTER 6
EMPIRICAL ASSESSMENT OF SYSTEMS-ORIENTED INSTRUCTION

6.1 Introduction

The interview methods described in the previous chapter were followed in two studies designed to assess systems-oriented instruction. The overall question guiding these efforts is as follows: Does learning in a systems-oriented curriculum change the way that students interpret information about complex social and ecological phenomena? In this chapter I will work through a number of more specific questions designed to develop an answer to the overall question. A description of these questions, the calculations designed to answer the questions, and the results from the calculations comprises the bulk of this chapter.

6.2 Description of Calculations

6.2.1 How to Read this Chapter

The purpose of this chapter is both to report specific results from two studies assessing systems-oriented instruction and to explore methods of analysis for such an assessment. Therefore, a high level of detail is used to describe the various calculations tested and reported for these studies. In this section I will provide an account of the calculations used in both studies and an explanation of why those calculations are performed. The subsections are organized by the questions that the each calculation is intended to address. The following two sections—covering the results from the two different studies using similar methodology—are organized with a structure parallel to this section. For example, subsection 2.2.1 explains calculations performed to compare those aspects of the study scenario that were deemed most important by each group. Then, subsections 3.2.1 and 4.2.1 provide the results of these calculations in each of the studies reported here. Organizing the chapter this way should provide the reader with a logical progression of information, but also facilitate moving back and forth between the calculations and the results from each study.

6.2.2 Analysis of 3CM Maps

Recall from the previous chapter that 3CM maps are developed to illustrate which aspects of system are deemed most important by the participants and how those aspects are associated. Each participant's 3CM map was converted to a Concept x Concept (in this case 36 x 36) matrix in which each cell ij shows whether or not concept i and concept j were included in the same group by the participant. Cells include a 1 to represent associated concepts or a 0 to show that concepts are not associated. Note that the diagonals of these matrices display which concepts were chosen, while cells off the diagonals display each participants' associations. These matrices from each group can be combined to form a similarity matrix in which each cell ij shows the percentage of the participants in that group who made an association between concepts i and j.

6.2.2.1 Was there a difference in the average number of concepts chosen by each group?

The number of concepts chosen by a participant is generally considered an indicator of the participants' confidence regarding his or her understanding of the topic being mapped (e.g. Amtmann 1996, Kearney *et al*. 1997). This figure was simply tallied during the data entry process. An analysis of variance was used to test for significance of the differences between the groups.

6.2.2.2 Was there a difference between groups regarding which concepts were chosen for the map?

The next step in the evaluation is simply to look at which concepts were chosen most often by each group and identify differences between groups. This question was addressed in two ways. First, a chi-squared test was used to identify significant differences in percentages of students from each group who chose each concept. Second, the concepts were aggregated into nine themes identified from a combination of hierarchical clustering and a content analysis of participant responses.[1] Table 6-1 lists the concepts by theme. The percentage of participants from each group who chose a concept from a particular theme was tallied. These percentages were compared across groups, just as the concept percentages were compared. The idea here is to get a sense of which themes received more attention from various groups.

6.2.2.3 Was there difference in how the participants from each group categorized their chosen concepts?

To address this question, the similarity matrix from each group was subjected to hierarchical clustering using SPSS. This process identifies stable categories of concepts chosen by participants within each group. These categories are then compared across groups.

6.2.3 Analysis of Cause Maps

Recall that participants produced cause maps during the second part of the mapping exercise by connecting nodes with arrows that illustrate the existence of a causal relationship and the direction of that relationship (i.e., whether an increase in the cause increases or decreases the effect). As with the 3CM maps, these maps were translated into 36 x 36 matrices. These are called adjacency matrices. Unlike the similarity matrices produced from the 3CM maps, adjacency matrices are directional. The rows are designated as causes and the columns as effects. Thus each cell ij contains information about the suggested link between cause i and effect j. A zero in cell ij shows that the participant drew no causal link from concept i to concept j; a 1 in cell ij shows that the participant drew a positive arrow from concept i to concept j, implying that an increase in concept i would produce an increase in concept j; and a (–1) in cell ij shows that the participant drew a negative arrow from concept i to concept j, implying an increase in concept i would produce a decrease in concept j.

Once the causal maps are translated into adjacency matrices various calculations can be made to provide information about the structure of each map and to compare maps. One potential problem with this type of analysis of causal maps is that it may fail to account for similar ideas on different maps. For example, Participant A may imply in their causal map that an increase of *Bad Weather* will decrease the *Effort put into catching samaki*, which will result in a decrease in the *Amount of Samaki Caught*. Participant B may express the same idea simply by showing that an increase of *Bad Weather* results in a decrease in the *Amount of Samaki Caught*. In this case, one might argue that the decrease in *Effort put into catching samaki* is implicit in Participant B's map. However, this similarity between Participant A and Participant B would not show up in a comparison of the adjacency matrices.

To address this problem, many of the calculations below are repeated with matrices that include indirect causal relationships. The adjacency matrices described above, showing only direct connections drawn by the participant, will be referred to as 1st-order matrices. By raising the 1st-order matrix to a power, one can identify indirect connections as well. For example, squaring a 1st-order matrix shows all of the two-step connections. Squaring Participant A's matrix in the example from the previous paragraph will show a connection from *Bad Weather* to *Amount of Samaki caught*, even those these concepts are not directly connected. Therefore, a matrix that included all direct connections and all two-step connections would be able to identify the similarity between Participant A's map and Participant B's map. I refer to such a matrix as a 2nd-order matrix.

Note that a 2nd-order matrix includes both direct and two-step connections. Similarly, a 3rd-order matrix includes direct, two-step, and three-step connections, and so on. In the measurements where the inclusion of indirect causal connections may make a significant difference, the calculation is made for all matrices up to the 5^{th-}order. Previous studies comparing causal maps (e.g. Ford and Hegarty 1984) have included all possible indirect connections in their calculations. With 36 available nodes, that would mean potentially including a 35-step connection.

Five was chosen as a reasonable limit in this study for a number of reasons. First, the interest here is in identifying those connections that the participants consciously included in their maps. It is unlikely that participants are consciously aware of connections more than five steps away. But more importantly, the inclusion of indirect causality is intended to identify similarities analogous to those between Participant A and Participant B above. A four-step chain is the longest chain that I could identify that could reasonably be contracted into one step. For example, consider the following chain: *Demand for farm-raised fish*→*Price per catch*→*Effort put into catching samaki*→*Amount of samaki caught*→*Samaki population*. Arguably, the intermediate steps here might be implicit in a map that shows an increase in *Demand in farm-raised fish* causing a decrease in the *Samaki population*. This similarity would be identified with fourth-degree matrices. Fifth-degree matrix calculations are included as a margin of safety.[2] In many cases, using higher-order matrices made little difference in the final result. However, since one purpose of this chapter is to explore the usefulness of available calculations, higher-order results are often reported.

6.2.3.1 Did the participants include web-like causality in their causal maps?

Recall from the previous chapter that Maani and Maharaj (2004) identify *operational thinking* as one of the three systems thinking skills that correlated most to better problem-solving performance. Barry Richmond's (2000) term, operational thinking, implies focusing on the causal structure of a system rather than simply identifying key drivers. We tend to think in linear causal chains that begin with key drivers because events occur in a sequence. However, the first step toward understanding dynamically complex systems is breaking out of the mindset that sees causality as occurring linearly and replacing the metaphor of a causal chain with that of a causal web. Therefore, the first analysis of participants' causal maps focuses on the inclusion of web-like causality.

There are three types of web-like causality. The first is simple branching as shown in Figure 6.1b, where nodes are connected by multiple branches rather than in a line. The second type may be called closed-branching. In this form, shown in Figure 6.1c, the branches close by linking back to other nodes. Closed-branching looks, at first glance, like the third type of web-like causality, causal loops. We will look more closely at causal loops in the next section. For now, the important thing to note is that even though the closed-branching looks like loops, one cannot get all the way around a closed-branch without going against the direction of at least one arrow.

The simplest way to measure web-like causality is link density, which is the ratio of links to node in a causal map. This measure is used for analyzing cognitive maps (Langfield-Smith 1992; Jenkins and Johnson 1997), as well as for analyzing ecological networks (Dunne et al. 2002). For example, in a 3-concept causal map depicting purely linear causality like Figure 6.1a, the link density would be 0.67 (2 links/3 nodes). However, if each concept affected every other concept, then there would be six links resulting in a link density of 2.0 (6 links/3 nodes). Link density will measure differences in closed-branching and causal loops, but does not measure changes simple branching. For example, Figure 6.2 shows two maps that each have a link density of 1.0 (7 links/7 nodes). But despite their equivalent link densities, they are very different maps for our purposes. Figure 6.2a shows an almost entirely linear map, while Figure

6.2b includes significantly more branching. Therefore, I have developed a second measure that is sensitive to changes in simple branching.

The second approach focuses specifically on the number of concepts that have more than one cause or more than one effect. In a completely linear map, no node would have more than one cause and one effect. Therefore, one can measure the degree to which a participant includes web-like causality by calculating for each map the percentage of nodes with more than one effect and the percentage of nodes with more than one cause. These two values are summed to produce a Web-like Causality Index (WCI). The larger the WCI score, the more a participant has stepped away from using purely linear causal structures. For example, in Figure 6.1, Map (a) has a WCI score of zero (0/3 + 0/3) and Map (b) has a WCI score of 0.33 (1/3 + 0/3). In Figure 6.2, Map (a) has a WCI score of 0.29 (1/7 + 1/7) and Map (b) has a WCI score of 0.57 (3/7 + 1/7). These two measurements in combination with the causal loop count—discussed below—can show the degree of web-like causality and provide information about the kind of web-like causality included.

6.2.3.2 Did the participants include causal loops in their causal maps?

We learned in Chapter 3 that positive and negative feedback loops are responsible for much of the counterintuitive behavior exhibited by dynamically complex systems. And you may recall from Chapter 5 that *closed-loop thinking*, the tendency to look for causal loops instead of linear causal chains, is a second systems thinking skill cited by Maani and Maharaj (2000) as corresponding to better participant performance. On a causal map, these loops are defined as a chain of causal connections that ends where it begins. A loop must contain at least two nodes. For example, the chain *Samaki population*→*Predatory bird population*→*Samaki population* implies a predator/prey relationship like that discussed in Chapters 2 and 3, in which each species affects the population of the other. Larger loops can also be included, such as *Amount of* *samaki caught*→*Samaki population*→*Public concern about over-fishing*→*Management of samaki population*→*Amount of samaki caught*. This loop represents the type of balancing feedback loop on which much government regulation of fisheries is based.

Despite the importance of feedback in understanding dynamically complex systems, studies—like those by Sterman (1989) and Moxnes (2000)—repeatedly illustrate individuals' lack of understanding of feedback loops. In his study on "naïve ideas about the causal order in natural processes," Peter White describes the causal chains identified by his study participants as "unremittingly linear" (1992, p. 70), explaining that out of 338 causal chains included in the participant responses, only one involved a causal loop. In this context, one might consider the inclusion of feedback loops in one's causal map an important step in applying systems thinking to understand a dynamically complex system. For this study, an ecological network analysis program called Windows Application for Network Analysis Digraphs (WAND) (Allesina 2000) was used to identify loops in the participant maps. The percent of participants from each group who included causal loops in their maps was then calculated and compared across groups.

6.2.3.3 Were there differences in the specific connections identified by participants from different groups?

Once we establish a difference in the level of connectivity included in maps by members of different groups, the next task is to identify which additional connections are being included. For these calculations an aggregate adjacency matrix was constructed for each group. These are 36 x 36 matrices in which cell ij shows the percentage of participants in a group who included a link from Concept i to Concept j on their causal map. Using these matrices we can identify which links were included more often by a particular group. We can then look at the sign (+ or -) on the link to interpret the relationship that the students intended. Aggregate adjacency matrixes were constructed using both 1st-order and 5th-order maps.[3]

In addition to looking at differences in specific connections, we can return to the nine themes described in Table 6-1. Here, we are looking at connections between themes instead of specific concepts. To do this, each participant's map is converted into a theme matrix—a 9 x 9 matrix where a 1 in cell ij implies that the participant included at least one connection from a cause in Theme i to at least one effect in Theme j. Then, a 9 x 9 aggregate theme matrix is constructed in which each cell ij shows the percentage of participants within a group who connected at least one cause in Theme i to at least one effect in Theme j.

6.2.3.4 How successful were the participants in accurately representing the issue described in the article?

The calculations so far, while characterizing how the participants understood the article, say little about the accuracy of their understanding. Indeed, 3CM maps are most helpful when there is no one right answer in mind, and while the causal map properties discussed here may exhibit a tendency toward looking for nonlinear causal structures, they do not necessarily imply better maps. In this section, I address the accuracy of the participants' causal maps.

While there is no one correct way to construct a cause map of the system described in the article, there are arguably key connections that one would need to make in order to demonstrate an understanding of the dynamics of the system. For example, a map that fails to show any connection between the demand for samaki and the samaki population may be said to be missing an important connection. In order to identify these key connections, a number of experts were asked to make their own maps of the situation described in the article. This resulted in four expert maps. The first two maps (labeled herein Exp1 and Exp2) were constructed by two groups consisting of both systems ecology professors and graduate students from the University of Florida. In each group, the participants read through the article individually and then worked together to choose applicable concepts and construct a causal map of the issue. The third and fourth maps (herein Exp3 and Exp4) were each constructed by individuals working alone—a systems modeler and a marine ecologist, respectively.

Two measurements were used to assess how close the participant maps are to the expert maps. First, a similarity index was calculated. When comparing any two maps, the similarity index is the ratio of the number of links common to both maps (mutual links) over the number of links unique to one map. This index was calculated to compare each participant's map with each of the four expert maps. A total score was then calculated for each participant by averaging the values from comparisons with each of the four experts. For example, Participant A's map was compared to Exp1, Exp2, Exp3, and Exp4, producing four similarity indices for Participant A. The average of these four indices is Participant A's total similarity score.

After all of the participants' total similarity scores were calculated, each group's mean score was calculated for comparison with other groups. Also, as a means for testing the internal validity of this measure, the similarity indices were calculated between the expert maps themselves. These calculations were repeated for 2nd-, 3rd-, 4th-, and 5th-degree causal maps.

One problem with a similarity index is that at very low scores it tends to reward those who included more connections. Therefore, an additional method of comparison was used. The second method for comparing participant maps to expert maps consisted of a rubric for scoring the participant maps. Each participant map was scored according to the rules in Table 6-2. The specific point values were dictated by two main concerns. First, the presumption is that a connection identified independently on all four expert maps counts as an important (or at least a clear) connection. Conversely, if all four experts have neglected to indicate a particular connection, it is likely that such a connection does not exist. Therefore, those connections receiving higher agreement among the experts are worth more points. Additionally, the numbers were chosen so that a purely random selection of connections would result, on average, in a slightly negative score.[4] The increased penalty on higher-order matrices for identifying connections that no expert identified maintains this slightly negative score for randomly selected connections. This ensures that participants will not be rewarded simply for choosing more connections.

Note that unlike the similarity index, those connections that were identified by the experts but not identified by a participant do not figure into this calculation. A participant is penalized only for identifying connections not identified by any expert. Therefore, a participant's rubric score can be interpreted as measure of the participant's success in identifying only those links that were also identified by the experts.

6.3 Study 1: Undergraduates from the University of Florida

6.3.1 Study Participants

The first study was a pre-test/post-test design, involving 23 undergraduate students in a political science class on environmental ethics and politics. Most of the students were political science majors, but other majors represented included English, journalism, marketing, and one environmental studies major. None of the students had received any systems training before taking this class.

While the class was not specifically devoted to systems concepts as with Lyneis and Lyneis (2003) and Pala and Vennix (2005), the material was presented from a systems perspective. That is, each new topic included explicit lessons on those concepts discussed in Chapter 3 and was explained in the context of those concepts. The course focused on understanding the nonlinear causal relationships within several natural resource issues, including population growth and family planning policies, water management, and global warming.

Participants worked through the mapping exercise once during the second week of the course and again at the end of the course, fourteen weeks later. The course did not cover any issue involving the management of a fishery. Therefore, whatever process or skills the students applied to the study problem during the post-test portion of the study had to be transferred from lessons given in other natural resource contexts.

Both the pre-test and post-test mapping exercises were a required part of the course. Both exercises were graded for completion, not for specific characteristics about their maps. Students were given the opportunity not to have their maps included as a part of the data for this study. Every student in the class filled out a form indicating their consent or refusal to have their coursework included as data in the study. These forms were returned in sealed envelopes, which were not opened until after final grades for the course were submitted. In this way, students could be confident that their decision would not affect their grade in the class.

6.3.2 Analysis of 3CM Maps

6.3.2.1 Was there a difference in the average number of concepts chosen by each group?

The participants chose an average of 16.0 concepts (SD = 3.6) during the pre-test and 20.1 concepts (SD = 5.5) during the post-test—a difference significant at $p < 0.003$. Recall that the number of concepts included is generally considered an indication of the participant's confidence in his or her knowledge about the issue. Even if this is the case, it is difficult to say here whether the increase in cards signifies an increased confidence with complex systems or simply greater comfort with natural resource issues having just completed a course on them.

6.3.2.2 Was there a difference between groups regarding which concepts were chosen for the map?

Table 6-3 shows all the concepts chosen by at least one third of the students in either the pre-test or the post-test, along with the percentage of students that chose each concept. A chi-squared test used to identify differences between individual concepts chosen during the pre-test and post-test shows that three concepts changed with a significance of at least $p < 0.05$. These were *Cost per unit catch, Nutrients in the water,* and *Amount of sport fish caught*. In all three cases, the students chose the concepts more often in the post-test than in the pre-test. However, it is difficult to draw any strong conclusions based on these three concepts.

Two other differences worth noting are in the concepts *Reproduction rate of samaki* and *Amount of samaki caught*. *Reproduction rate of samaki* was chosen by 17% (4 students) during the pre-test and 43% (10 students) during the post, and *Amount of samaki caught* was chosen by 65% (15 students) in the pre-test and 96% (22 students) in the post. These concepts are particularly important in terms of systems thinking because they represent an in-flow and out-flow for the samaki population, which suggests a dynamic view of samaki with population levels changing due to other aspects in the system. While these differences are not statistically significant, they are the first signs of a trend we observe of a more nuanced view of the samaki population.

6.3.2.3 Was there a difference in how the participants from each group categorized their chosen concepts?

The hierarchical cluster analysis produced four distinct categories in the pre-test and five distinct categories in the post-test. Table 6-4 shows the stable categories identified for the pre- and post-tests. For each test, one of the clusters constituted a sort of miscellaneous category. Concepts from this category are not listed. As before, only those concepts chosen by at least one-third of the students are included.

Notice that the first three clusters, regarding demand, economic aspects, and ecological aspects, remain relatively constant from pre-test to post-test.[5] The fourth cluster, not included in the pre-test results, shows *Samaki population* and *Amount of samaki caught* as a separate category. This extra category, may again suggest a more nuanced view of the resource as its own entity, affecting and being affected by both the second and third clusters.[6]

6.3.3 Analysis of Causal Maps

6.3.3.1 Did the participants include web-like causality in their causal maps?

The mean link density was 1.11 for the pre-test maps and 1.32 for the post-test maps, meaning that participants increased their link density by an average of 0.21 from pre-test to post-test, a statistically significant difference at 95% confidence. This difference is significant in a practical sense as well. Since the nodes chosen by participants increased for the post-test, the increase in link density achieved here requires a mean increase of almost nine links on each post-test map.

The mean Web-like Causality Index (WCI) score for the pre-test and post-test results are 0.48 and 0.62, respectively, meaning that the students increased their WCI scores by an average of 0.13, a significant difference at 95% confidence levels. The results for both the link density and the WCI suggest that the students did indeed include more web-like causality in their post-test maps.

6.3.3.2 Did the participants include causal loops in their causal maps?

Six students (26%) included at least one causal loop in their maps during the pre-test. Five of those six students included at least one loop in their post-tests as well. (One who included a causal loop in his pre-test did not include one in the post-test.) In addition, seven students who did not include a causal loop in their pre-tests included at least one loop in their post-tests. This makes a total of twelve students (52%) who included causal feedback in their post-test maps. The sample size is too small for this to be a statistically significant change, and these results are hardly overwhelming considering the foundational position of causal loops in understanding dynamic complexity.

But looking more closely at the types of loops included by the students gives a brighter picture. Most of the students (four out of six), who included loops in the pre-test, included only two-node loops. That is, their sense of feedback seems restricted to a two-node relationship in which Node 1 affects Node 2 and Node 2 in turn affects Node 1. While this type of two-way causality is important, it is arguably the easiest type of loop to identify. In the post-test, conversely, nine students included loops involving three or more nodes. This provides some evidence that these nine students have a deeper sense of feedback loops than they did in the pre-test. Nonetheless, the failure of almost half of the students to include any loops at all suggests that more needs to be done to stress the importance of feedback.

6.3.3.3 Were there differences in the specific connections identified by participants from different groups?

Using a chi-squared analysis, ten first-order causal links were identified that were chosen significantly more often during the post-test than the pre-test. These pairs are shown in Table 6-5. The calculations were then repeated using 5th-order matrices. As expected, including indirect links produced much different results. Fifty causal links were identified as significantly more common in the post-test than the pre-test.[7] Of these fifty links, five match with the results from the 1st-order calculations. Table 6-6 shows the results for these five links along with other notable differences.

The first five rows of Table 6-6 display the connections that were significantly different with both 1st-order and 5th-order calculations. The other connections included on the table illustrate the greater connection between catch levels and fish populations and the greater role that *Management of samaki catch* plays in the post-test results. Again, since the course itself focused on natural resource issues, the greater role of resource management cannot be attributed solely to systems-oriented instruction. However, the greater focus on population flows is an important part of systems thinking.

In this case, the 9 x 9 theme matrices did not add any insights. Only one theme-to-theme connection (from Weather to Industry) was identified as changing significantly ($p < 0.05$) from pre-test to post-test. This connection was made by 17% of the participants in the pre-test and 61% in the post-test. The lack of significant difference in any other theme-to-theme connection suggests that the additional connections in the post-test were spread out across the themes.

6.3.3.4 How successful were the participants in accurately representing the issue described in the article?

Table 6-7 summarizes the results from the Total Similarity Score calculations. Since this study involves dependent samples, the difference between each participant's pre-test and post-test performances was used as the test statistic. Table 6-7 shows the Total Similarity Scores from the pre-test and post-test results and the mean difference between pre-test and post-test results, including upper and lower limits for each difference calculated at 95% confidence levels. One can see by the positive value of both the upper and lower limits that the post-test results had higher similarity scores (i.e., were more similar to expert maps) than the pre-test results regardless of which order matrices were used for the calculation. However, the difference becomes more pronounced as more indirect links are included in the calculation. This trend can be attributed to the higher-order calculations picking up similar ideas expressed in slightly different ways as described at the beginning of Section 2.3. The right column of Table 6-7 shows the Pearson correlation between each calculation and the fifth-order calculation. The high correlation between fourth- and fifth-order calculations suggests that the fourth-order calculations would likely suffice in picking up indirect causal similarities between maps.

As a basis for comparison, Table 6-8 shows Total Similarity Scores between the four expert maps. These scores were calculated by comparing each expert map to the other three expert maps. With four experts, there are six possible pairings.[8] The mean score listed in Table 6-8 represents the mean of these six similarity scores. The maximum and minimum similarity scores for any two expert maps are also included. The point here is not to provide a statistical analysis, but simply to provide values to which one can compare the similarity scores achieved by the students. As expected, the expert maps are much more similar to each other than are the student maps to the expert maps.

Comparisons based on the rubric in Table 6-9 yields a picture similar to Table 6-7. Table 6-9 shows that improvement from the pre-test to the post-test is significant at 99% confidence limits regardless of which order matrices are used in the calculation. Recall that the rubric is designed to slightly favor smaller maps when random connections are chosen. Still, improvement from pre-test to post-test becomes more pronounced as more indirect causal connections are included, with the exception of the 5^{th}-order calculations which slow slightly less improvement than is shown in the 4^{th}-order calculations. Again, the high correlation between 4^{th}-order and 5^{th}-order calculations suggests that little is gained by extending calculations to the 5^{th}-order.

6.3.4 Summary and Discussion

6.3.4.1 Relative Focus of Concepts and Connections

To review, the students chose significantly more cards during the post-test than the pre-test, suggesting a greater confidence with the content of the article. Both *Reproduction rate of samaki* and *Amount of samaki caught* were chosen more often during the post-test. While these differences were not statistically significant, they foreshadowed differences we would see again in the hierarchical clustering analysis, in which a Samaki population cluster suggested a more nuanced view of samaki as a dynamic population playing a significant role in the system.

Analysis of the specific causal connections showed a continuation of this trend with more attention being paid during the post-test to the relationship between the samaki population and the samaki catch-rates.[9] A related trend also shows up here concerning the greater role that management plays regarding both the samaki and sport fish populations. And finally, we see in the post-test connections from environmental aspects (e.g. *Coastal water quality*) that would likely increase the perceived need for some sort of environmental management. These findings suggest a more dynamic view of the samaki population, a clearer sense of its role in the larger system, and a greater understanding of the role that management plays in the system.

6.3.4.2 Structural Differences and Comparison to Expert Maps

The students included significantly more web-like causality during the post-test. While the moderate increase in the use of causal loops was somewhat disappointing, it still showed marked improvement in the understanding of the role that causal loops play in dynamically complex systems. The post-test maps also showed greater similarity to the four expert maps, even when measured by the rubric that penalizes students for including connections not included by the experts. These findings suggest that the students were better able at the end of the class to understand the complex environmental system described in the article.

6.3.4.3 Student Opinions

My extended contact with these students allowed me to acquire information about how the students themselves feel about systems concepts. These opinions were solicited in two ways. First, after completing the post-test, students were given both their pre-test and post-test maps and asked to compare the two, identifying and explaining key differences. These opinions were not anonymous and were collected before the end of the course, so the positive views expressed about systems concepts may reflect a student's interest in his course grade more than his true opinion of the course content. Nonetheless, the insights shared by some of the students are encouraging.[10]

For example, one student observes a shift of focus that corresponds to the findings from the hierarchical cluster analysis—"For the first map I put 'Disagreement over Samaki Population' as the primary card, which in the second map I put 'Samaki population'." Another student expresses dissatisfaction with her pre-test map because of its lack of feedback: "everything in the first map seems to all filter in one direction: toward the samaki population card (as inputs usually negatively affecting it)." A third student explains why she thinks the increased connections on her post-test map reflect a more accurate picture of the system: "For example, 'Management of Samaki Fish' had only one arrow drawn into it on the first map and none going out; while on my second map it had three going out, because I did not take into account the effect of this change on the rest of the system in the first map. I can now see how many things would be effected by this change." These statements—and others like them—suggest that whether or not students are truly convinced of the value of systems thinking, they have at least learned to describe its potential value.

The second method for soliciting student opinions addresses some of the shortcomings of the first. At the end of the course each student filled out a survey about their self-assessed understanding of systems concepts and the likelihood that they would apply these concepts in the future. These surveys were returned in sealed envelopes that were not opened until after the course grades had been submitted.

The results of this survey are reported in Table 6-10. Note that the students received a more complete discussion of systems concepts than was provided in Chapter 3, so the language in the survey is slightly different. The students were asked to assess their understanding of four systems concepts: web-like causality, stocks and flows, scale, and adaptive cycles (i.e., a model for how complex systems change over time). Table 6-10 shows the mean response scores with 1 representing no understanding at all and 5 representing a very strong understanding. Students were then asked about the likelihood that they would use these concepts in the future when assessing environmental issues. The mean scores for these questions are reported with 1 representing no chance and 5 representing a very high likelihood. The last two questions pertain to systems concepts in general.

On the questions about understanding, the mean scores are all between "Fair understanding" (3) and "Above average understanding" (4). And for the likelihood of using these concepts in the future, the scores are all between "Somewhat likely" (3) and "Likely" (4). Scores for all the concepts except "adaptive cycles" were far closer to the "Above average understanding" and "Likely future use" marks. These responses suggest that the students themselves recognize the usefulness of systems thinking.

6.3.4.4 Shortcomings of the Study

While the data here suggest that students, after receiving systems-oriented instruction, were better able to comprehend information about a complex environmental issue, there are a number of factors—besides systems-oriented instruction—that could have affected the results. First, as mentioned several times, the course focused on natural resource issues. Thus, any change from pre-test to post-test might be the result of a semester of being forced to learn and think about these issues. Second, because of the course's explicit focus on systems concepts, the students would likely have been conscious of the types of causal structures to look for during the post-test. Addressing these shortcomings requires, first, performing this assessment in a course focused on something other than natural resources and, second, separating the assessment from the course itself so that the exercise does not overtly cue students to use systems concepts. One could then be more confident that any use of systems concepts reflected a more general use of systems thinking skills by the student. These differences are included in the following study.

6.4 Study 2: Portland Study

6.4.1 Description of Participants

6.4.1.1 Comparison of study groups

The second study involved 147 middle school students from Portland Public School system in Portland, Oregon. These students represent four groups—two groups of seventh graders and two groups of eighth graders. One group from each grade came from a school (herein called School A) in which the vast majority of teachers incorporate systems-oriented instruction into their curricula. These participants are particularly appropriate for this kind of study because, unlike the participants from Study 1, most of them have been exposed to systems concepts for more than one year. The two control groups were comprised of students from a school (herein called School B) in the same district as School A, but do not use systems-oriented instruction.

For all four groups, the mapping exercise was performed as a classroom activity. Consent forms were sent home to the students' parents giving them the option to not have their students participate in the exercise. Those students who did participate did not receive a grade or any other compensation for their participation.

None of the four groups had received specific training in environmental issues. The systems concepts used in School A were not presented in the context of natural resources. Therefore, any application of systems thinking exhibited by the students can be interpreted as evidence that they apply systems thinking on their own, without prompting. Also, unlike the first study, none of these participants were aware that the evaluation of the maps would focus on systems thinking skills. The students were told only that I was interested in how they interpret the information contained in the study article.

Table 6-11 shows how the two schools compare in terms of profile and performance on standardized tests.[11] Since data for the 2005-2006 school year is not available at the time of write-up, the data for Table 6-11 is from the 2004-2005 school year. During the 04-05 school year, the study participants were in sixth and seventh grade. Therefore, data for those grades are included below. One can see that while School B's scores are high, School A's scores are exceptionally so—roughly 15% higher than those of School B. Additionally, School A had a significantly higher percentage of students exceeding, rather than meeting, the benchmarks. These differences are partly explained by the fact that, in addition to using systems-oriented instruction, School A has an accelerated curriculum and attracts students interested in such a curriculum. This difference was unavoidable, and School B is considered by local district representatives to be the best comparison to School A available.

Nonetheless, the difference in performance presents a problem in terms of direct comparison. How can one be sure that any differences between the performance of School A participants and School B participants in the present study are due to systems-oriented instruction rather than the generally high performance illustrated by the standardized test scores? I address this problem in two ways. First, while school B does not have tracking at the seventh-grade level, it does have tracking at the eighth-grade level. The eighth-grade participants from School B were taken from a pool of students in the accelerated mathematics track. Tracking was not an issue at School A, as all of the students (except those absent on the days of testing) were included in the study. Therefore, while standardized test performance was not available from individual students, it is likely that the standardized test results from the eighth-grade participants from School B would more closely resemble those from School A than would a random sample of eighth graders from School B. A second way of addressing School A's exceptional standardized test performance and accelerated curriculum is to compare School A's seventh graders with School B's eighth graders. These comparisons are included throughout the discussion.

6.4.1.2 Systems curriculum at School A

The philosophy of introducing systems-oriented instruction into the curriculum at School A is a grassroots one. That is, no teacher is required to use systems-oriented instruction at all, much less to use it in a specific way. The rationale behind this approach is that teachers who are forced to implement a particular pedagogical tool or method are less likely to implement that tool effectively. The method for diffusing this tool at School A is to offer assistance for teachers wanting to implement the tool in their teaching and to provide professional certification credit for courses taught about systems-oriented instruction. Over time, enough teachers from School A have become convinced of the usefulness of the tool that now even when new teachers come to School A, they become curious about the tool themselves. As a result, without top-down regulation, every teacher at School A implements systems-oriented instruction in some way. However, the extent to which each teacher implements systems-oriented instruction varies widely.

The instruction is guided by fostering twelve "habits of a systems thinker" as outlined by the Waters Foundation, a non-profit organization promoting and developing systems-based instruction. According to the Waters Foundation, a systems thinker:

1. Seeks to understand the "big picture."
2. Observes how accumulations within systems change over time.
3. Changes perspectives to increase understanding.
4. Identifies the circular nature of complex cause and effect relationships, i.e., interdependencies.

5. Understands and considers how mental models affect current reality and the future
6. Considers both short and long term consequences of actions.
7. Surfaces and tests assumptions.
8. Finds where unintended consequences emerge.
9. Recognizes that a system's structure generates its behavior; focuses on structure, not on blame.
10. Uses understanding of system structures to identify possible leverage actions.
11. Holds the tension when considering issues without trying to resolve it quickly.
12. Checks results and changes actions if needed: successive approximation. (Waters Foundation 2003)[12]

These habits are conveyed in a number of difference ways. In the lower grades the teachers often introduce systems concepts in the context of learning classroom behavior. For example, one teacher leads her students through an exercise in which they identify connections between various classroom rules. Other teachers employ systems-based instruction to help students understand the events in a novel by drawing connections between the events to identify causal structures. Comprehension of novels as well as historical events is enhanced through a tool called the iceberg model (Waters Foundation 2003) where students learn to see events as the tip of the iceberg and are encouraged to look beneath the surface to identify patterns of behavior, underlying causal structures, and mental models that have resulted in the observed events. One teacher has used this tool extensively to her lessons on ancient history to point out parallels between various ancient societies. Note that in all of these examples the content of the lessons does not differ greatly from other schools in the district. However, the students at School A are taught to think about this content in a fundamentally different way from students learning in conventional curricula.

Despite the variations between teachers regarding the use of systems-based instruction, School A represents one of the most well established programs for systems-based instruction in the country. Also, the variation in teachers' use of systems-oriented instruction does not introduce variability within the School A study groups. School A has only one teacher per grade at the lower grades (K-5th) and at the upper grades (6th-8th) students split their time between two teachers. In other words, all students attending School A at the same grade level receive the same instruction.

6.4.2 Analysis of 3CM Maps

6.4.2.1 Was there a difference in the average number of concepts chosen by each group?

Herein, the study groups will be designated as A7, B7, A8, and B8, where the letter refers to groups from School A or School B and the number refers to the grade of the students. Groups A7 and B7 both chose an average of 11.0 concepts (SD = 4.1 and SD = 4.4 respectively). Since this measure is considered to indicate participants' confidence in their understanding of the topic, this result suggests that the two groups have a similar level of confidence in their ability to understand the information in the article. This result is not surprising because both groups have had very little instruction in natural resources or ecology.

However, the eighth-graders showed much different results. Group A8 chose an average of 15.1 concepts (SD = 5.9), while Group B8 chose an average of 10.0 (SD = 4.3). This difference is significant with a p-value < 0.001. Group A8's results seem to be in line with those from the seventh graders. One would expect eighth graders to display more confidence in the exercise than the younger students. However, Group B8's average is actually slightly lower than Group B7's, but with low significance ($p < 0.4$). This relative lack of confidence on the part of B8 is puzzling, but as we shall see, it is the first sign of a trend in B8 to perform similarly to B7.

6.4.2.2 Was there a difference between groups regarding which concepts were chosen for the map?

Before looking at differences between the groups, it may be helpful to get a feel for which concepts were chosen most often in general. Table 6-12 shows all the concepts chosen by at least 40% of the students in any of the four groups along with the percentage of students that chose each concept. As one might expect Samaki population was chosen most often by all four groups.

A chi-squared test used to identify differences between individual concepts chosen by the two seventh-grade groups yields only two concepts chosen at significantly different percentages ($p < 0.05$). These were *Dissolved oxygen levels*, which 4% of A7 chose compared to 19% of B7, and *Sales price/unit catch (samaki)*, which 14% of A7 chose compared to 37% of B7. Due to the relatively low percentage for *Dissolved oxygen levels* from both parties, it does not play a notable role in other evaluations. However, the higher percentage from B7 of *Sales price/unit catch (samaki)* foreshadows a greater focus (and perhaps understanding) by B7 on economic aspects of the issue.

With the eighth-grade groups, five concepts were chosen at significantly higher percentages by A8 than by B8 ($p < 0.05$). These concepts are listed in Table 6-13. Again, the differences here foreshadow more significant differences later. First, there is the inclusion of *Reproduction rate of samaki*. As explained in the previous study, in order to understand a dynamically complex system, one must identify the important stocks and the in-flows and out-flows associated with these stocks. Reproduction of samaki represents the in-flow to the population of samaki.

The second, third, and fourth concepts in Table 6-13 pertain to the role that public perception plays in the issue. We will see later that students from A8 are able to connect matters of public perception to both demand and environmental management. The inclusion of *Dissolved oxygen levels*, as with the seventh-graders, does not represent a significant part of later calculations.

When one looks at the chosen concepts in terms of the nine themes from Table 6-1, no significant differences arise between the two seventh-grade groups or between the two eighth-grade groups. This means that at each grade level each of the nine themes was identified by roughly the same percentage of people from the systems group and the corresponding control group. The totals for all four groups are shown in Table 6-14.

6.4.2.3 Was there difference in how the participants from each group categorized their chosen concepts?

Table 6-15 shows the groups that result from the hierarchical clustering analysis of the two seventh-grade similarity matrices. For each group, one of the clusters constituted a sort of miscellaneous category. Concepts from this cluster are not listed. As before, only those concepts chosen by at least 40% of the participants in at least one group are included.

The first three clusters, regarding samaki population, samaki catch, and demand, are identical. However, the fourth and fifth clusters show notable differences. The systems group, A7, has a much larger cluster regarding the various ecological aspects of the samaki issue. In contrast, the control group, B7, does not put these concepts together. This may suggest a greater sense on the part of the systems group of how these concepts relate to one another. However, the situation is reversed in the fifth cluster. Here, the control group puts together a number of concepts dealing with various economic aspects relating to the resource. This category does not exist at all for the systems group.

Table 6-16 shows the hierarchical clustering results for the eighth graders. As with the seventh graders, one of the clusters for each group constituted a sort of miscellaneous category. Concepts from this cluster are not listed. Again, only those concepts chosen by at least 40% of the participants from at least one group are included.

Both groups identify a category for demand and nearly identical categories for ecological aspects. In addition, both groups identify a category on samaki population, but an interesting difference arises here. The systems group includes *Amount of samaki caught* and *Reproduction rate of samaki* in this category. In other words, in addition to identifying the samaki population as a stock, the systems students group this concept with the in-flow (reproduction) and out-flow (samaki catch) for the stock. The grouping of samaki stocks and flows in one category suggests a dynamic (or systems) understanding of the resource. And finally, the systems group includes a fourth category regarding economic aspects that the control group does not include. This is the converse of the seventh-grade groups in which the control group was the only one to include this category. However, unlike Group B7, Group A8 also includes a well-developed Ecological Aspects group

6.4.3 Analysis of Causal Maps

6.4.3.1 Did the participants include web-like causality in their causal maps?

Table 6-17 shows the students' mean scores for both link density and the Web-like Causality Index (WCI). Group A7 has a significantly higher link density than B7 and a significantly higher WCI score. These results suggest that A7 includes more web-like causality than the control group, regardless of whether the calculation includes simple branching or not.

Group A8 shows a link density that is significantly higher than both B8 ($p < 0.005$) and A7 ($p < 0.005$). A8's WCI score is also significantly higher than both B8 ($p < 0.005$) and A7 ($p < 0.025$).[13] This implies that A8 has significantly more web-like causality regardless of whether the calculation includes simple branching or not. Comparing the control groups shows that B8 has slightly higher link density and B7 has a higher WCI score. However, neither of these differences is statistically significant. And finally, comparing A7 to B8 shows that the seventh-grade systems group includes significantly more web-like causality as measured by link density ($p < 0.05$) and WCI ($p < 0.05$).

6.4.3.2 Did the participants include causal loops in their causal maps?

Among the seventh graders, two students (7.4%) from the control group and eighteen students (37%) from the systems group included at least one causal loop in their maps. This implies a significant difference ($p < 0.025$) between the two groups' inclusion of causal feedback. The systems students also included more complicated feedback loops. Of the two students from the control group who included feedback, one student included one two-node loop and the other student included one three-node loop. In the systems group twelve students included more than one causal loop and eleven students included loops with three or more nodes.

Among the eighth graders, seven students (23%) from the control group and nineteen students (48%) from the systems group included at least one causal loop. This difference is significant at $p < 0.1$. Again, this difference appears more significant when one looks more closely at the inclusion of feedback. In the systems group, 11 students (28%) include causal loops involving three or more nodes and 14 students (35%) include more than one causal loop. In the control group, five students (16%) include a loop of three or more nodes and three students (10%) include more than one loop. Comparing Group A7 with Group B8 shows that a higher proportion of the systems seventh graders include causal loops, but this difference is not statistically significant at $p < 0.1$. Nor is the difference between the two systems groups, A7 and A8, or between the two controls, B7 and B8.

6.4.3.3 Were there differences in the specific connections identified by participants from different groups?

Comparison of the 1^{st}-order aggregate adjacency matrices from the seventh-grade groups shows three causal links that are identified more commonly by A7. These are listed in Table 6-

18, along with the percentage of students from each group that identified the links. Among these three connections, the first is the most significant in terms of systems thinking. Over half of A7 connects the stock population to an out-flow related to that stock. The second link in Table 6-18 is an important relationship, but the percentage difference is not great enough to draw any strong conclusions. And the third link in the table implies an intermediate *Samaki population* node, so its inclusion in this 1st-order analysis might be considered of little importance. However, this link remains significant in the 5th-order aggregate adjacency matrix as shown in Table 6-19.

A total of fourteen links are chosen at significantly different percentages in the 5th-order calculation. Notably, the link from *Amount of samaki caught* to *Samaki population* also remains significant. This, along with the other connections involving *Samaki population,* again suggests a greater emphasis on the connection between the samaki population, out-flow from the population (fish catch), and a number of other factors, including management and coastal water quality. Also, over a fifth of A7 identify a feedback loop involving the samaki population.

Six of the fourteen significant differences resulted from B7 picking connections more frequently than A7. Of these, links 11 and 12 from Table 6-19 continue the trend of the seventh-grade control group to identify more economic connections than their systems counterparts. However, other connections included by the control group are more difficult to interpret. Link 9 (*Donestre & Sons' profits*→*Algae blooms/Dead zones*) is perhaps feasible with several intermediate concepts (e.g. *Donestre & Sons' profits*→*Effort put into catching samaki*→*Amount of samaki caught*→*Samaki population*→*Algae blooms/ Dead zones*), but other connections chosen by B7 are questionable. For example, it is difficult to envision within the context of the article how the *Amount of samaki caught* could affect the *Demand for farm-raised fish,* barring an aggressive advertising campaign on the part of Donestre & Sons. I will suggest a reason for this and other puzzling links in the discussion below.

Looking at the 1st-order connections in terms of the nine themes from Table 6-1 we can see four connections chosen at significantly different percentages ($p < 0.05$). These connections are shown in Table 6-20. In the first connection, A7 suggests that demand for various samaki products also affects *Production of international fish-oil producers*, *Prices of competing products (e.g. soybeans)*, and *Soybean sales*. And with the second connection, they suggest how public concern regarding the samaki fishery can lead to increased efforts to better manage the resource. Both of these connections seem reasonable.

The second two connections, chosen more frequently by B7, are somewhat more difficult to interpret. Potentially, if a decrease in water quality was attached in the public mind to demand for products that require samaki, then demand for those products could go down, but since *Public worry about decrease of samaki* is not chosen as an intermediate link, it is difficult to know if this is the series of events the students had in mind. The fourth connection stems from B7 students suggesting that *Bad Weather* will affect predatory bird, marine mammal, and sport fish populations. Again, this connection is difficult to interpret. Potentially, *Bad Weather* could mean less samaki fishing by humans and, therefore, more samaki left for other predators to eat. But again, the *Samaki population* was not chosen as a likely intermediate concept. In both of these cases, the connections were made by a relatively small number of students from B7. Still, they reflect a trend on the part of B7 students to identify confusing or nonsensical connections at significantly higher percentages.

The trends that begin to appear among the seventh graders are seen more clearly with the eighth graders. A chi-squared test identifies twenty causal links that are chosen at significantly higher percentages in A8 than in B8 in the 1st-order aggregate causal adjacency matrices. These links are listed in Table 6-21. The first four links suggest that demand for products made with samaki will affect the amount of samaki caught. Links 5 through 10 then lay out a number of aspects that depend on the amount of samaki caught. Links 11 and 12 continue this causal web to include effects on other species, as do links 13 and 14. However, 13 and 14 are particularly interesting in that they express the type of positive feedback loop described in Chapter 3 that defines basic population dynamics. The remaining six links include the dynamics of public debate over the issue and secondary economic factors.

Only three of the above differences remain significant in the 5th-order aggregate adjacency matrix. Table 6-22 lists the thirty-one causal links that show a significant difference between Groups A8 and B8 ($p < 0.05$). The first nine connections deal with the samaki catch levels and the effects it has on both ecological and social systems, including Link 2 to the *Samaki population* itself. Links 10 through 13 continue this theme with the *Samaki population* as the cause concept. Links 14 through 18 deal with other social and ecological aspects that affect samaki, including resource management and public concern, Links 19 through 23 pertain to the connections between public perception and resource management, and Links 24 through 27 cover other miscellaneous direct and indirect links.

The last four links represent the four connections made more commonly by B8 than by A8. Link 28 shows the most direct connection, suggesting that an increase in the effort required to catch samaki would increase the costs. Links 29 and 30 make sense if one equates an increase in *Donestre & Sons' profits* with a decrease in the *Samaki population*. But Link 31, which as seen in B7 as well, is difficult to interpret.

And finally, looking at the theme-to-theme connections from the eighth graders, we can see seven theme-to-theme connections chosen at significantly different percentages ($p < 0.05$). These are listed in Table 6-23. The first link shows a wide difference in the percentage of students from each group identifying a link from the samaki industry to the samaki population, and Link 2 shows almost as wide a difference in the connection from the samaki population to water quality issues. Links 3 through 5 all pertain to the connections between the samaki industry, public perception, and environmental management. The last two connections were identified more often by B8. Link 6 can be interpreted as the suggestion that a decrease in water quality will effect either public perception or the samaki population itself in such a way that will affect the samaki industry. But Link 7, suggesting that the samaki somehow affect the weather, is another connection that is difficult to interpret.

The general trend in this analysis of concept and theme connections seems to be an abundance of often important connections chosen by a significantly higher percentage of students from A7 and A8. In addition, several of the links chosen by a significantly higher percentage of students from B7 and B8 are difficult to interpret sensibly. One explanation for these connections could be a misunderstanding of the mapping process. However, I offer another possible explanation in the discussion below.

6.4.3.4 How successful were the participants in accurately representing the issue described in the article?

Table 6-24 summarizes the results from the similarity index calculations for the seventh graders, including the mean total similarity scores for each group and the significance of the difference between those means. The seventh-grade systems group shows significantly greater similarity to the expert maps than the control group regardless of which order matrices are used to perform the calculations. Scores for both groups increase slightly with higher-order calculations, with the exception of the fifth-order score for Group A7. The high correlation between 3^{rd}-, 4^{th}-, and 5^{th}-order calculations suggests that little is gained in the analysis by using 4^{th}- and 5^{th}-order connections.

Table 6-25 summarizes the same data for the eighth-grade groups.[14] The difference here is even more pronounced with the eighth-grade systems group more than doubling the similarity scores of the eighth-grade control group regardless of which order matrices are used for the calculations. Again, the correlation between 3^{rd}-, 4^{th}-, and 5^{th}-order results suggests that using 3^{rd}-order matrices here would have sufficed. [15]

Comparison between the two systems groups, A7 and A8, shows that the eighth-graders produce significantly higher similarity scores ($p < 0.025$) regardless of which order matrices are used for the calculations. This result may suggest that students receiving systems-oriented instruction continue to improve their ability to understand complexity in eighth grade. Conversely, comparison between the two control groups shows that the eighth-grade control group did not produce higher similarity scores than their seventh-grade colleagues. In fact, the seventh grade control group produced slightly higher similarity scores, although none of the differences between the control groups are significant within a 95% confidence limit. Also worth noting, comparison between A7 and B8 shows that the seventh graders receiving systems-oriented instruction produce significantly higher similarity scores than the eighth-grade control group ($p < 0.015$) regardless of which order matrices are used in the calculations.

These same relationships hold for the rubric data as well. Table 6-26 shows the mean rubric scores for the seventh-grade groups, tallied according to the rules outlined in Table 6-2. As with the similarity scores, the systems group convincingly outperforms the control group regardless of which order matrices are used in the calculation. For both groups the rubric scores continue to rise for higher-order calculations. This result is expected because the rubric does not penalize participants for missing the extra indirect links contained in the expert maps. The more important result is that the control group's score remains consistently at a little less than half the systems group's score. As with the similarity scores, the correlation between 4^{th}- and 5^{th}-order calculations is very high, suggesting that little is gained by the 5^{th}-order calcultions.

The rubric results from the eighth-graders, shown in Table 6-27, follow a similar pattern, but again, the differences are more pronounced. The control group's score is roughly one quarter that of the systems group's regardless of which order matrices is used in the calculation.[16] The 3^{rd}- and 4^{th}-order calculations are both highly correlated with the 5^{th}-order calculations.

Comparison between the two systems groups shows that Group A8 outperforms A7 at 95% confidence levels for 1^{st}- and 2^{nd}-order calculations only. The lack of significance at higher levels can be explained by the increased variability in the higher-order results. Table 6-28 shows this trend, using A8 results as an example. The reason for this trend is that higher-order rubric calculations can exaggerate both correct and incorrect connections (as judged against expert agreement).

Using the extreme cases as examples should make this point clearer. Table 6-29 shows the results from two participants from the A8 group. Notice that while Participant A8-38 outperforms Participant A8-16 at every level of calculation, the difference changes greatly. Two different trends are occurring. In the case of A8-38, more of the participant's connections match the experts' picks as more indirect connections are included. In other words, misses (connections A8-38 chose that no expert chose) begin to be scored as hits at higher-order calculations. This trend is, in fact, the reason for investigating higher-order calculations.

Alternatively, while A8-16 scores above his group's mean score in the 1st-order calculation, he falls further and further below the mean as more indirect connections are included. In this case, A8-16 has chosen a few misses involving nodes that are highly connected. These misses result in a fairly minor penalty in the 1st-order calculation, but they increase quickly as more indirect connections involving those misses are included and as the penalty for misses becomes more severe. This trend is problematic because it creates highly volatile scores. For example, removing one connection from A8-16's map changes his higher-order scores dramatically, as shown in the right column of Table 6-29. It should be noted that this pattern of going from above the mean in the 1st-order calculation to below zero in higher order calculations occurs only one other time out of the 147 participants. Nonetheless, the potential for this volatility, combined with the relatively high correspondence between lower-order calculations and the 5th-order calculations may be reason enough to restrict rubric calculations to lower-order matrices.

6.4.4 Summary and Discussion

6.4.4.1 Relative Focus of Concepts and Connections

To review the Portland study, we started the 3CM analysis by looking at the number of concepts chosen by each group. There was no significant difference between the mean number chosen by the seventh-grade groups, and Group B8 averaged roughly the same number as the two seventh-grade groups. A8, however, averaged significantly higher than the other three groups. These results suggested a greater confidence on the part of A8 in their understanding of the article. Also, the fact that B8 did not exhibit greater confidence than the seventh-grade groups was the first sign that B8's performance would be on par with that of B7.

Next we looked for differences in the concepts chosen and found that among the seventh-graders *Sales price/unit catch* was chosen more often by B7 than by A7. This was the first sign of a greater emphasis on economic aspects on the part of B7. The trend continued in the clustering analysis, where B7 had a fully developed economic cluster and A7 did not. However, the clustering analysis also showed that A7 had a more fully developed cluster containing ecological connections. This trend continued in the analysis of specific causal connections. A7 included several ecological connections at significantly higher percentages than B7, and B7 included two economic relationships that A7 did not include at all. Thus, we might characterize the relative focus of A7 as leaning to the ecological aspects of the issue and that of B7 as leaning toward the economic aspects.

With respect to the eighth-graders, A8 chose *Reproduction rate of samaki* significantly more often than B8. Reproduction of samaki is an in-flow to the samaki population. Therefore, its inclusion suggests that A8 is more likely to look at samaki as a stock that changes depending on its in-flows and out-flows.

The hierarchical clustering analysis of the eighth-grade 3CM maps showed that, unlike the seventh-graders, it was the systems group, A8, that included a well-developed economic cluster not included by the control group. However, in this case A8 also included a well-developed cluster for ecological connections, suggesting an awareness of ecological and economic aspects. Analysis of the individual causal connections supported this finding. A8 identified at a higher percentage than B8 a large number of connections from demand to samaki catch, from samaki catch to samaki population, from samaki population to other ecological processes, from public perception to environmental management, and from environmental management to the samaki population. Analysis of theme-to-theme connections produced similar differences. This suggests a relatively high understanding by member of A8 of both economic and ecological connections.

6.4.4.2 The Role of Public Perception

Looking at the theme-to-theme causal connections for the seventh-graders showed another trend concerning environmental management. Eighteen percent of A7 participants identified a connection from Public Perception to Environmental Management compared to zero percent of B7. This difference showed up again with the eighth graders. A8 included three concepts dealing with public perceptions more often than B8. This attention to public perception continued in the causal analysis with A8 including several more indirect connections pertaining to public perceptions than did B8. And in theme-to-theme connections, A8 chose both Public Perception to Environmental Management and Public Perception to Industry at significantly higher percentages than B8. The inclusion of public perception is particularly significant because of the role it plays in feedback loops pertaining to the management of the resource.

6.4.4.3 Structural Differences between Maps

Structurally, both systems groups A7 and A8 showed significantly more web-like causality and a significantly higher percentage of students including feedback loops than did their respective control groups. In addition, A7 showed significantly more inclusion of web-like causality than did B8. This is partly because the B8 did not include more web-like causality than B7. This finding seems to suggest not only that students receiving systems-oriented instruction are more likely to use non-linear structures to interpret non-linear systems, but also that without systems-oriented instruction, students are not likely to improve in their abilities to understand dynamically complex systems. The fact that A8 recorded a higher mean link density and WCI score than that recorded in the pre-test of the undergraduates from Study 1 provides more evidence for this trend. This finding supports the studies cited in Chapters 3 and 4 that suggest that individuals are not natural systems thinkers (e.g. Moxnes 2000, Sterman and Booth Sweeney 2002).

6.4.4.4 Linear Thinking and Unlikely Connections

Another trend that we can see in the analysis of the cause maps is the identification on the part of the control groups of connections that are difficult to interpret sensibly, such as *Amount of samaki caught* → *Demand for Omega-3*. Some of these connections may be due to a lack of understanding of the exercise itself. For example, *Demand for Omega-3*→*Public information to increase fish-oil intake*, which was identified by both B7 and B8 at significantly higher percentages than their respective treatment groups, is a nonsensical connection. However, reversing the direction of the arrow yields a highly plausible relationship. Conceivably, the presence of this connection in the control groups is simply a misunderstanding of the arrow syntax.

But in that other unlikely connections cannot be so easily explained by simply reversing the direction of the causal arrow, a second possibility may go further in explaining these odd connections. Some background may be appropriate here. During the pilot tests for this research too, I worked with a number of individuals representing a broad range of ages. During one of these pilot-test sessions, I was working with a particularly vocal eighth-grader who wanted very much to come up with the "right" answer to the causal mapping exercise. As he worked, he placed his chosen cards in a line and quickly became frustrated with the lack of causal options offered by his linear model. After asking me a number of questions about where the connections should go and receiving only evasive answers ("Put what you think is best."), he shifted to thinking aloud and studying my facial expressions for signs that he was on the right track. As a result, I got the chance to hear his thought process as he worked through the exercise.

The student had made a chain of roughly four nodes, but was having difficulty placing a fifth node at the end of the chain. He would repeatedly try a card, reason aloud why that node was inappropriate, and take it away again. This continued until he had worked through all of his chosen cards and not found a suitable solution. Finally, he began adding cards and lengthening his causal chain with which he was audibly unsatisfied, but he saw them as the best he could come up with. The possibility of deviating from a purely linear structure did not occur to him despite his concerted effort. The maps in Figure 6.3 taken from B7 and B8 results remind me of this student's frustration. I suggest that the peculiar links included in these maps (and others like them) are not due to a misunderstanding of the meaning of the arrows, but rather the same kind of frustration that student felt during the pilot test.[17]

6.4.4.5 Comparison to Experts

Both systems groups developed causal maps that were, on average, more similar to expert maps than were their respective control groups. In addition the maps produced by A7 were more similar to expert maps than were maps produced by B8. These findings suggest that students receiving systems-oriented instruction are better able to interpret information about dynamically complex social and ecological systems. And finally, A8 produced causal maps more similar to the expert maps than the undergraduates from Study 1 produced during the pre-test, supporting the hypothesis that conventional curricula does not teach students to better understand complexity.[18]

6.5 Conclusions

6.5.1 Lessons Learned Regarding Calculations

A second line of questioning involved the feasibility of the various indices and calculations used to evaluate the causal maps. While the WCI score presented here did not produce results different from the link density index used in other studies, the reasons for its inclusions still stand. In the present studies, differences in web-like causality were achieved through closed-branching and causal loops, making the link densities significantly different between groups.

However, if the web-like causality included consisted primarily of simple branching, then the WCI score would be necessary to identify the difference.

Regarding the inclusion of indirect causal connections, the 5^{th}-degree calculations seemed to offer little over the 4^{th}-degree calculations. And with the rubric scores used to compare participant maps to expert ones, the higher-order calculations produced the potential for volatile results. The present studies seem to suggest that 4^{th}-order calculations will suffice in identifying similarities between maps' indirect connections. However, repeated studies would be necessary to test this limit.

6.5.2 Group Comparisons

Recall that the overall question guiding these two studies was whether systems-oriented instruction changed the way that students interpret information about a complex social and ecological system. Results from both studies seem to suggest that this is indeed the case. With the first study, involving undergraduates from the University of Florida, we saw that in the post-test the students exhibited a more nuanced view of the samaki population and its role in the larger system, as well as more links involving environmental management. In addition, post-test maps included more web-like causality and scored more similarly to the expert maps. However, given that these students were learning systems concepts in the context of natural resource issues, one might attribute the performance during the post-test to the course in general and not specifically to the systems-oriented approach. The Portland study addresses this concern.

Unlike the study involving undergraduates, these students were not learning specifically about natural resources. Nonetheless, the two systems groups seem to be better prepared to understand the complex environmental issue presented to them in this study. That is, the systems groups included more web-like causality (including more causal loops), identified more key causal connections (including those involving public perceptions), and scored more similarly to expert causal maps than their respective control groups.

Table 6-30 summarizes the results of both studies and includes figures from the four expert maps to facilitate comparisons between all four participant group and the expert maps. Note that that Group A8 in the Portland study outperforms the undergraduates during the pre-test in the U.F. study. If one accepts the assumption that sophomores and juniors at the University of Florida are likely to be better readers than eighth-graders from Portland, then these results suggest that this tool measures something other than reading comprehension—specifically, ability to understand complex relationships. However, this tool could be used in conjunction with a tool measuring reading comprehension to verify this hypothesis. Note also that, as one would expect, the experts included more weblike causality than any of the participant groups, and the expert maps were more similar to each other than to any of the participant groups' maps.

As these studies involve less than 200 students, one must be cautious about extrapolating the results. Still, the results do suggest that systems-oriented instruction may be an appropriate pedagogical tool for helping students comprehend dynamically complex systems. Certainly, these results are promising enough to warrant further studies that replicate the structure outlined here.

Notes

[1] Details of the results from the hierarchical clustering will be discussed below.

[2] While this is not a formal modeling exercise, we might also heed the warning from modeler Galomb: "Don't believe the 33rd-order consequences of a first-order model" (1968, qtd. in Odum 1983, p.580).

[3] While the signs (+ or -) on the arrows help to characterize the connection, they do not reflect the magnitude of a connection and are, therefore, not suitable for analyses involving indirect causal links. For this reason, they are used as aids for understanding the 1st-order connections.

[4] For example, in the first order map, only four connections show up on all four expert maps, 25 connections show up on three maps, 39 on two maps, and 113 on only one. With these proportions, a random selection of a connections will average out to –0.009 per connection chosen.

[5] The cluster labels were added by the researcher for presentation purposes.

[6] The position of the samaki population as its own category is also supported by the post-test association between the *Samaki population* and *Management of samaki catch*. While these two concepts are not associated at all with each other in the pre-test, 30% of the participants associated these two concepts in the post-test

[7] To test the feasibility of these extra post-test links, they were compared to the four expert maps. Five of these fifty extra links did not appear on any of the expert maps, suggesting that they are perhaps unnecessary or infeasible. However, the other forty-five connections (including all the connections reported in Table 6.6) were identified on at least one of the expert maps.

[8] The possible pairings are as follows: 1-2, 1-3, 1-4, 2-3, 2-4, 3-4.

[9] More attention was also paid to the relationship between sport fish populations and their catch-rates.

[10] This exercise was conducted more as pedagogical tool than a data collecting one.

[11] These data are taken from reports issued by Portland Public Schools.

[12] For a more detailed discussion of these habits, see Booth Sweeney's and Meadows' *Systems Thinking Playbook* (2001).

[13] Notice that A8's scores here are higher than the undergraduates' scores from the pre-test in Study 1.

[14] Refer to Table 6-7 to compare these scores to the similarity scores between the four expert maps.

[15] First-order pre-test and post-test total similarity scores from Study 1 are 0.179 and 0.278 respectively. Thus, the systems eighth graders produced causal maps more similar to expert maps than the pre-test maps produced by the undergraduates from Study 1.

[16] First-order pre-test and post-test rubric scores from Study 1are 21.9 and 37.0 respectively. As with the similarity index, these scores suggest that the systems eighth graders produced causal maps more similar to expert maps than the pre-test maps produced by the undergraduates from Study 1.

[17] This frustration is not unlike that felt by Moxnes' participants who believed that their poor performance in the reindeer-lichen simulation must have been due to causes outside the framework of the problem (e.g. computer error)

[18] This also suggests that the assessment tool used here is not simply a reading comprehension test.

Table 6-1. Concept cards listed according to themes

Theme	Concept
Environmental Management……..	Management of samaki catch
	Management at ecosystem level
Ecological connections…………...	Marine mammal population
	Predatory bird population
	Sport fish health
	Sport fish populations
	Amount of sport fish caught
Water Quality…………………….	Algae blooms/ Dead zones
	Coastal water quality
	Dissolved oxygen levels
	Nutrients in the water
Commercial Demand……………..	Demand for farm-raised fish
	Demand for livestock
	Demand for Omega-3 food supplement
	Public information to increase Omega-3 intake
Samaki Industry…………………..	Amount of samaki caught
	Cost per unit catch
	Sales price per unit catch
	Effort put into catching samaki
	Donestre & Sons' Profits
Samaki Population………………..	Samaki population
	Lifespan of samaki
	Reproduction rate of samaki
	Reproduction rate per unit fish
	Food eaten per unit fish
Weather…………………………	Bad weather
	El Nino
Public Perception of Environment..	Public worry over decrease of samaki
	Scientific speculation of overfishing
	Donestre & Son's claim of a healthy fishery
	Disagreement over samaki population health
Indirect Economic Aspects……….	Society affluence
	Soybean sales
	Price of competing products (e.g. soybeans, vegetable oils)
	Production from international fish-oil competitors
	Human population

Table 6-2. Rules for scoring participant maps against expert agreement

Identifying a connection in agreement with:	Four experts	8 points
	Three experts	6 points
	Two experts	4 points
	One expert	1 point
	Zero experts (First-order)	-0.5 points
	Zero experts (Second-order)	-2 points
	Zero experts (Third-order)	-4 points
	Zero experts (Fourth-order)	-6 points
	Zero experts (Fifth-order)	-8 points

Table 6-3. Concepts chosen during pre- and post-tests

Concept	Pre-test %	Post-test %
Algae blooms/ Dead zones	91	100
Samaki population	91	96
Demand for Omega-3 food supplement	87	100
Demand for farm-raised fish	70	96
Demand for livestock feed	78	87
Donestre & Sons' profits	70	91
Amount of samaki caught	65	96
Scientific speculation of over-fishing	78	70
Price of competing products (soybeans and veg. oils)	74	65
Public information to increase fish oil intake	61	78
Coastal water quality	61	70
Management of samaki catch	57	74
Bad weather	48	70
Disagreement over health of samaki fishery	74	43
Sport fish health	43	70
Soybean sales	48	61
Sport fish populations	48	61
Production from international fish oil competitors	52	52
Donestre & Sons' claim of a healthy fishery	52	48
Management at the ecosystem level	48	52
Public worry about decrease of samaki	43	52
El Nino	43	48
Predatory bird populations	39	52
Nutrients in the water	22	61
Sales price per unit catch	22	43
Society affluence	13	52
Reproduction rate of samaki	17	43
Cost per unit catch	4	52

Table 6-4. Stable categories chosen by participants

Category 1	Category 2	Category 3	Category 4
Pre-Test			
• Demand for farm-raised fish	• Donestre & Sons' profits	• Algae blooms/ Dead zones	
• Demand for livestock feed	• Price of competing products	• Coastal water quality	
• Demand for Omega-3	• Production from international competitors	• Sport fish populations	
• Public information about fish oil intake	• Soybean sales	• Sport fish health	
	• Sales price per unit catch	• Predatory bird populations	
		• Samaki population	
Post-Test			
• Demand for farm-raised fish	• Donestre & Sons' profits	• Algae blooms/ Dead zones	• Amount of samaki caught
• Demand for livestock feed	• Price of competing products	• Coastal water quality	• Samaki population
• Demand for Omega-3	• Production of international competitors	• Sport fish populations	
• Public information about fish oil intake	• Sales price per unit catch (samaki)	• Sport fish health	
	• Soybean sales	• Predatory bird populations	
	• Cost per unit catch (samaki)	• Amount of sport fish caught	
	• Bad Weather	• Nutrients in the water	
	• El Nino		

Table 6-5. Percentage of participants identifying causal links in 1st-order matrices among undergraduates

	Cause	Effect	Pre-Test %	Post-Test %
1	Algae blooms/ Dead zones	Sport fish populations	0	22
2	Bad weather	Donestre & Sons' Profits	4	35
3	Cost/unit catch (samaki)	Donestre & Sons' Profits	0	17
4	Cost/unit catch (samaki)	Amount of samaki caught	0	20
5	Sport fish populations	Amount of sport fish caught	4	30
6	Demand for Omega-3	Donestre & Sons' Profits	9	48
7	Samaki population	Donestre & Sons' Profits	0	22
8	El Nino	Donestre & Sons' Profits	9	22
9	Algae blooms/ Dead zones	Scientific speculation of over-fishing	0	20
10	Management at Ecosystem Level	Management of samaki catch	0	20

Table 6-6. Percentage of participants identifying causal links in 5th-order matrices among undergraduates

	Cause	Effect	Pre-Test %	Post-Test %
1	Algae blooms/ Dead zones	Sport fish populations	4	30
2	Bad weather	Donestre & Sons' Profits	17	52
3	Cost/unit catch (samaki)	Donestre & Sons' Profits	0	17
4	Cost/unit catch (samaki)	Amount of samaki caught	0	20
5	Amount of samaki caught	Reproduction rate of samaki	0	20
6	Amount of samaki caught	Samaki population	35	78
7	Management of samaki catch	Amount of samaki caught	0	30
8	Management at ecosystem level	Sport fish populations	0	17
9	Coastal water quality	Management of samaki catch	0	20
10	Predatory bird populations	Management of samaki catch	0	20
11	Samaki population	Amount of sport fish caught	4	35
12	Sport fish populations	Amount of sport fish caught	4	39

Table 6-7. Comparisons of total similarity scores for pre- and post-tests

	Total Similarity Scores		Mean Difference	Upper and Lower Limits of Mean Difference (95% Confidence)		Pearson Correlation to 5th-Order ($p < 0.01$)
	Pre-Test Mean	Post-Test Mean				
1st Order	0.045	0.070	0.025	0.010	0.040	0.621
2nd Order	0.043	0.085	0.041	0.021	0.062	0.816
3rd Order	0.046	0.103	0.057	0.027	0.086	0.887
4th Order	0.050	0.117	0.066	0.031	0.102	0.915
5th Order	0.054	0.126	0.072	0.033	0.111	1.0

Table 6-8. Total similarity scores between expert maps

	Mean	Minimum	Maximum
1st Order	0.237	0.110	0.380
2nd Order	0.290	0.107	0.538
3rd Order	0.454	0.146	0.843
4th Order	0.587	0.193	1.23
5th Order	0.623	0.189	1.53

Table 6-9. Comparisons of rubric scores for pre- and post-tests

	Rubric Scores		Mean Difference	Upper and Lower Limits of Mean Difference (99% Confidence)		Pearson Correlation to 5th-Order ($p < 0.01$)
	Pre-Test Mean	Post-Test Mean				
1st Order	21.9	37.0	15.1	2.71	27.4	0.542
2nd Order	42.7	107	64.3	19.6	109	0.825
3rd Order	60.9	184	123	27.3	218	0.898
4th Order	73.3	242	168	39.2	297	0.956
5th Order	69.7	209	139	37.7	241	1.0

Table 6-10. Survey responses regarding systems concepts

	Question	Score
1.	How would you characterize your level of understanding of web-like causality?	3.82
2.	How would you characterize your level of understanding of stocks and flows?	3.86
3.	How would you characterize your understanding of scale?	3.89
4.	How would you characterize your understanding of adaptive cycles?	3.39
5.	When you are considering current events regarding environmental issues, how likely are you to consider weblike causality?	3.71
6.	When you are considering current events regarding environmental issues, how likely are you to consider stock and flow?	3.82
7.	When you are considering current events regarding environmental issues, how likely are you to consider scale?	3.75
8.	When you are considering current events regarding environmental issues, how likely are you to consider adaptive cycles?	3.29
9.	How would you characterize your knowledge of systems concepts in general?	3.75
10.	When you are considering current events regarding environmental issues, how likely are you to consider systems concepts in general?	3.89

Table 6-11. Basic statistics of schools included in the study

2005 Statistics	School A	School B
Enrollment (Grades 6-8)	161	496
Budget Per Student	$4,501	$4,710
% that Met or Exceeded Benchmarks on Standardized Tests		
6^{th} Grade Reading	98%	90%
6^{th} Grade Math	98%	82%
7^{th} Grade Reading	98%	86%
7^{th} Grade Math	100%	85%

Table 6-12. Concepts included in cognitive maps by each group

Concept	A7%	B7%	A8%	B8%
Samaki population	90	89	95	85
Amount of samaki caught	90	78	88	56
Demand for Omega-3 food supplement	59	56	83	53
Algae blooms/ Dead zones	57	56	73	47
Public worry about decrease of samaki	57	48	55	29
Donestre & Sons' profits	41	48	60	41
Reproduction rate of samaki	45	37	48	12
Predatory bird populations	41	33	45	35
Demand for farm-raised fish	33	41	60	47
Effort put into catching samaki	29	41	20	29
Cost per unit catch (samaki)	24	44	35	15
Demand for livestock feed	37	33	80	56
Coastal water quality	33	11	53	32
Disagreement over samaki population health	27	30	58	26
Scientific speculation of over-fishing	27	33	53	18
Public information to increase fish oil intake	27	33	55	15
Price of competing products (soybeans and veg. oils)	16	11	43	21
Soybean sales	24	22	43	18

Table 6-13. Differences in the percentages at which concepts were chosen by eighth-grade groups

	Concepts	A8 (%)	B8 (%)
1	Reproduction rate of samaki	48	12
2	Scientific speculation of over-fishing	53	18
3	Disagreement over health of samaki fishery	58	26
4	Public information to increase fish-oil intake	55	15
5	Dissolved oxygen levels	18	0

Table 6-14. Percentage of students from each group that chose concepts for particular themes

	Theme	A7 (%)	B7 (%)	A8 (%)	B8 (%)
1	Environmental Management	41	44	45	26
2	Ecological connections	69	56	73	74
3	Water Quality	63	59	80	65
4	Commercial Demand	74	93	93	81
5	Samaki Industry	96	96	95	84
6	Samaki Population	94	100	95	97
7	Weather	33	26	40	52
8	Public Perception of Environment	71	67	80	55
9	Indirect Economic Aspects	49	41	65	52

Table 6-15. Hierarchical clustering results for seventh-grade participants

Category 1	Category 2	Category 3	Category 4	Category 5
Group A7				
• Samaki population	• Amount of samaki caught	• Demand for farm-raised fish • Demand for livestock feed • Demand for Omega-3	• Algae Blooms/ Dead Zones • Coastal water quality • Nutrients in the water • Predatory bird populations • Marine mammal populations • Sport fish populations	
Group B7				
• Samaki population	• Amount of samaki caught	• Demand for farm-raised fish • Demand for livestock feed • Demand for Omega-3	• Algae Blooms/ Dead Zones	• Cost/unit catch (samaki) • Effort put into catching samaki • Donestre & Sons' profits • Sales price/unit catch (samaki) • Food eaten/fish (samaki)

Table 6-16. Hierarchical clustering results for eighth-grade participants

Category 1	Category 2	Category 3	Category 4
Group A8			
• Samaki population	• Demand for farm-raised fish	• Algae Blooms/ Dead Zones	• Donestre & Sons' profits
• Amount samaki caught	• Demand for livestock feed	• Coastal water quality	• Sales price/unit catch (samaki)
• Reproduction rate of samaki	• Demand for Omega-3	• Predatory bird populations	• Cost/unit catch (samaki)
		• Marine mammal populations	• Soybean sales
		• Sport fish populations	• Public information to increase fish oil intake
		• Sport fish health	• Price of competing products (e.g. soybeans)
		• Nutrients in the water	• Production of international competitors
Group B8			
• Samaki population	• Demand for farm-raised fish	• Algae Blooms/ Dead Zones	
	• Demand for livestock feed	• Coastal water quality	
	• Demand for Omega-3	• Predatory bird populations	
		• Marine mammal populations	
		• Sport fish populations	
		• Sport fish health	
		• Amount of sport fish caught	

Table 6-17. Web-like causality mean scores

Group	Link Density	WCI Score
A7	1.10	0.46
B7	0.892	0.28
Significance of Difference	$p < 0.01$	$P < 0.01$
A8	1.29	0.59
B8	0.958	0.24
Significance of Difference	$p < 0.005$	$P < 0.005$

Table 6-18. Percentage of participants identifying causal links in 1st-order matrices among seventh graders

	Cause	Effect	A7%	B7%
1	Amount of samaki caught	Samaki population	53	22
2	Public information to increase fish-oil intake	Demand for Omega-3	10	0
3	Amount of samaki caught	Algae blooms/ Dead zones	16	0

Table 6-19. Percentage of participants identifying causal links in 5th-order matrices among seventh graders

	Cause	Effect	A7 (%)	B7 (%)
1	Amount of samaki caught	Algae blooms/ Dead zones	29	4
2	Amount of samaki caught	Samaki population	53	22
3	Management of samaki catch	Samaki population	14	0
4	Samaki population	Coastal water quality	20	0
5	Samaki population	Nutrients in the water	10	0
6	Samaki population	Samaki population	22	4
7	Sales price/unit catch	Management of samaki catch	20	0
8	Bad weather	Amount of samaki caught	14	0
9	Donestre & Sons' profits	Algae blooms/ Dead zones	0	11
10	Amount of samaki caught	Demand for farm-raised fish	0	20
11	Effort put into catching samaki	Sales price/unit catch (samaki)	0	19
12	Donestre & Sons' profits	Sales price/unit catch (samaki)	0	15
13	Demand for Omega-3	Public info to increase fish-oil intake	0	10
14	Samaki population	Lifespan of samaki	0	10

Table 6-20. Differences in theme connections among seventh graders

	Cause	Effect	A7(%)	B7 (%)
1	Demands	Indirect Social Factors	22	4
2	Public Perception	Ecological Management	18	0
3	Water Quality	Demands	0	11
4	Weather	Ecological connections	2	15

Table 6-21. Percentage of participants identifying causal links in 1st-order matrices among eighth-graders

	Cause	Effect	A7 (%)	B7 (%)
1	Public info to increase fish-oil intake	Demand for Omega-3	30	0
2	Demand for farm-raised fish	Amount of samaki caught	23	3
3	Demand for livestock feed	Amount of samaki caught	28	6
4	Demand for Omega-3	Amount of samaki caught	28	6
5	Amount of samaki caught	Samaki population	53	6
6	Amount of samaki caught	Algae blooms/ Dead zones	15	3
7	Amount of samaki caught	Donestre & Sons' profits	30	0
8	Amount of samaki caught	Public worry over decrease of samaki	10	0
9	Amount of samaki caught	Scientific speculation of over-fishing	20	0
10	Amount of samaki caught	Cost/unit catch (samaki)	10	0
11	Samaki population	Algae blooms/ Dead zones	35	6
12	Samaki population	Predatory bird populations	20	3
13	Samaki population	Reproduction rate of samaki	10	0
14	Reproduction rate of samaki	Samaki population	30	0
15	Scientific speculation of over-fishing	Management of samaki catch	10	0
16	Scientific speculation of over-fishing	Public worry over decrease of samaki	20	0
17	Scientific speculation of over-fishing	Disagreement over samaki population health	20	0
18	Donestre & Sons' claim of a healthy fishery	Disagreement over samaki population health	10	0
19	Sales price/unit catch (samaki)	Donestre & Sons' profits	10	0
20	Soybean sales	Donestre & Sons' profits	20	0

Table 6-22. Percentage of participants identifying causal links in 5th-order matrices among eighth-graders

	Cause	Effect	A7 (%)	B7 (%)
1	Bad weather	Amount of samaki caught	18	0
2	Amount of samaki caught	Samaki population	65	19
3	Amount of samaki caught	Algae blooms/ Dead zones	35	3
4	Amount of samaki caught	Donestre & Sons' profits	28	6
5	Amount of samaki caught	Predatory bird populations	23	3
6	Amount of samaki caught	Donestre & Sons' claim of healthy fishery	10	0
7	Amount of samaki caught	Coastal water quality	10	0
8	Amount of samaki caught	Amount of samaki caught	25	3
9	Effort put into catching samaki	Algae blooms/ Dead zones	15	0
10	Samaki population	Samaki population	28	3
11	Samaki population	Coastal water quality	20	0
12	Samaki population	Donestre & Sons' claim of healthy fishery	20	0
13	Samaki population	Nutrients in the water	20	0
14	Marine mammal populations	Samaki population	13	0
15	Effort put into catching samaki	Samaki population	23	3
16	El Nino	Samaki population	13	0
17	Management of samaki catch	Samaki population	18	0
18	Public worry about decrease in samaki population	Samaki population	28	6
19	Public worry about decrease in samaki population	Management of samaki catch	20	0
20	Public worry about decrease in samaki population	Public worry about decrease in samaki population	15	0
21	Public worry about decrease in samaki population	Donestre & Sons' claim of healthy fishery	20	0
22	Disagreement over samaki population health	Management of samaki catch	10	0
23	Management of samaki catch	Public worry about decrease in samaki population	13	0
24	Management of samaki catch	Marine mammal populations	13	0
25	Human population	Algae blooms/ Dead zones	13	0
26	Lifespan of menhaden	Management of samaki catch	10	0
27	Coastal water quality	Algae blooms/ Dead zones	15	0
28	Effort put into catching samaki	Cost/unit catch samaki	0	16
29	Donestre & Sons' profit	Algae blooms/ Dead zones	0	10
30	Donestre & Sons' profit	Cost/unit catch (samaki)	0	13
31	Demand for Omega-3	Public info to increase fish-oil intake	0	10

Table 6-23. Differences in theme connections among eighth graders

	Cause	Effect	A8 (%)	B8 (%)
1	Industry	Samaki	65	16
2	Samaki	Water quality	45	16
3	Industry	Public Perception	25	3
4	Public perception	Environmental Management	20	3
5	Public perception	Industry	28	7
6	Water Quality	Industry	3	16
7	Samaki	Weather	0	10

Table 6-24. Comparison of Seventh-Grade Total Similarity scores*

	Total Similarity Scores			Pearson Correlation to 5th-Order ($p < 0.01$)
	A7 Mean	B7 Mean	Significance of Difference	
1st Order	0.151	0.0966	P < 0.011	0.615
2nd Order	0.150	0.0912	P < 0.002	0.856
3rd Order	0.164	0.0982	P < 0.008	0.954
4th Order	0.172	0.101	P < 0.016	0.992
5th Order	0.169	0.102	P < 0.027	1.0

Table 6-25. Comparison of eighth-grade total similarity scores

	Total Similarity Scores			Pearson Correlation to 5th-Order ($p < 0.01$)
	A8 Mean	B8 Mean	Significance of Difference	
1st Order	0.222	0.0751	P < 0.001	0.788
2nd Order	0.235	0.0902	P < 0.001	0.916
3rd Order	0.256	0.101	P < 0.001	0.979
4th Order	0.260	0.103	P < 0.001	0.993
5th Order	0.267	0.102	P < 0.001	1.0

Table 6-26. Comparison of seventh-grade rubric scores

	Rubric Scores			Pearson Correlation to 5^{th}-Order ($p < 0.01$)
	A7	B7	Significance of Difference	
1^{st} Order	17.4	9.13	0.007	0.527
2^{nd} Order	41.4	21.4	0.003	0.713
3^{rd} Order	69.4	34.4	0.003	0.869
4^{th} Order	92.8	43.0	0.006	0.933
5^{th} Order	108	53.8	0.019	1.0

Table 6-27. Comparison of eighth-grade rubric scores

	Rubric Scores			Pearson Correlation to 5^{th}-Order ($p < 0.01$)
	A8	B8	Significance of Difference	
1^{st} Order	27.5	6.61	0.007	0.706
2^{nd} Order	69.5	20.9	0.003	0.826
3^{rd} Order	103	31.6	0.003	0.909
4^{th} Order	116	30.8	0.006	0.948
5^{th} Order	147	32.0	0.019	1.0

Table 6-28. A8 results with increasing variability at higher-orders

	Mean	Minimum	Maximum	Standard Deviation
1^{st}-Order	27.5	-3.5	112	22.1
2^{nd}-Order	69.5	-7	239	56.9
3^{rd}-Order	103	-34	366	94.8
4^{th}-Order	116	-174	449	123
5^{th}-Order	146	-282	803	178

Table 6.29. Scores from two members of group A8

	A8-16	A8-38	A8-16 (Altered)
1^{st}-Order	41	69	41.5
2^{nd}-Order	64	188	84
3^{rd}-Order	-34	366	46
4^{th}-Order	-174	449	-3
5^{th}-Order	-282	803	-48

Table 6.30. Summary of results from both studies and experts

Group	Nodes	Link/Node	WCI	% Including Feedback Loops	Total Similarity Score (4th Order)	Rubric Score (4th Order)
UF Pre-Test	16.1	1.11	0.483	26%	0.050	73.3
UF Post-Test	20.1	1.32	0.616	52%	0.117	242
A7	11.0	1.10	0.458	37%	0.172	92.8
B7	11.0	0.892	0.285	7.4%	0.101	43.0
A8	15.1	1.29	0.591	48%	0.260	116
B8	10.0	0.958	0.239	23%	0.103	30.8
Experts	31.5	2.16	1.08	100%	0.587	N/A

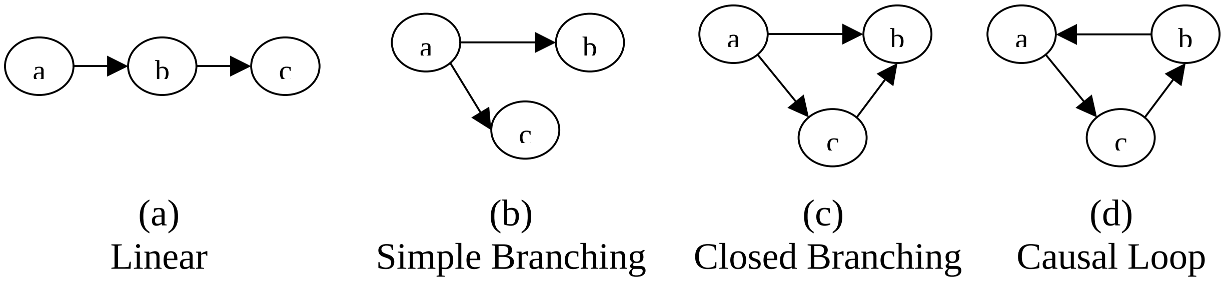


Figure 6-1. Types of web-like causality

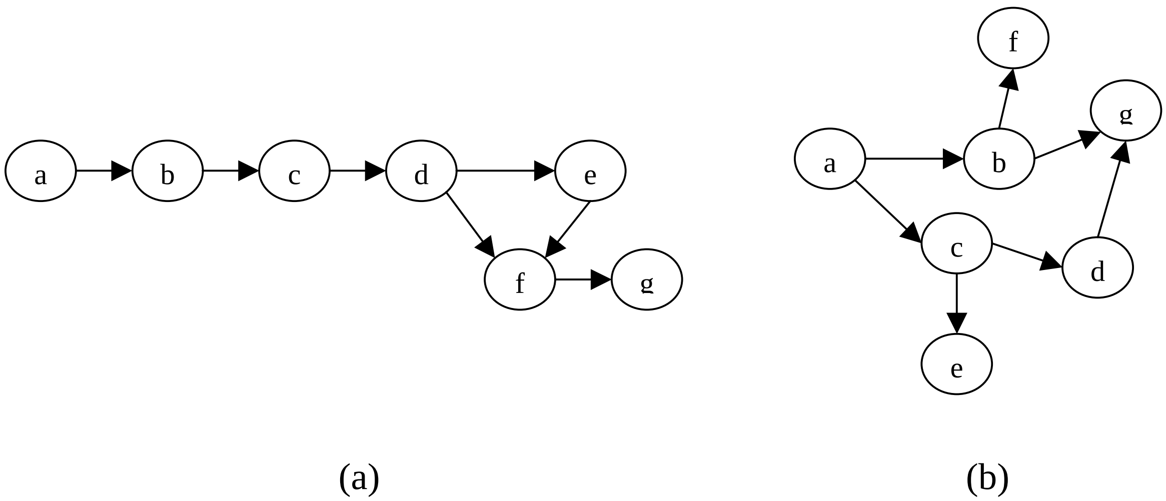
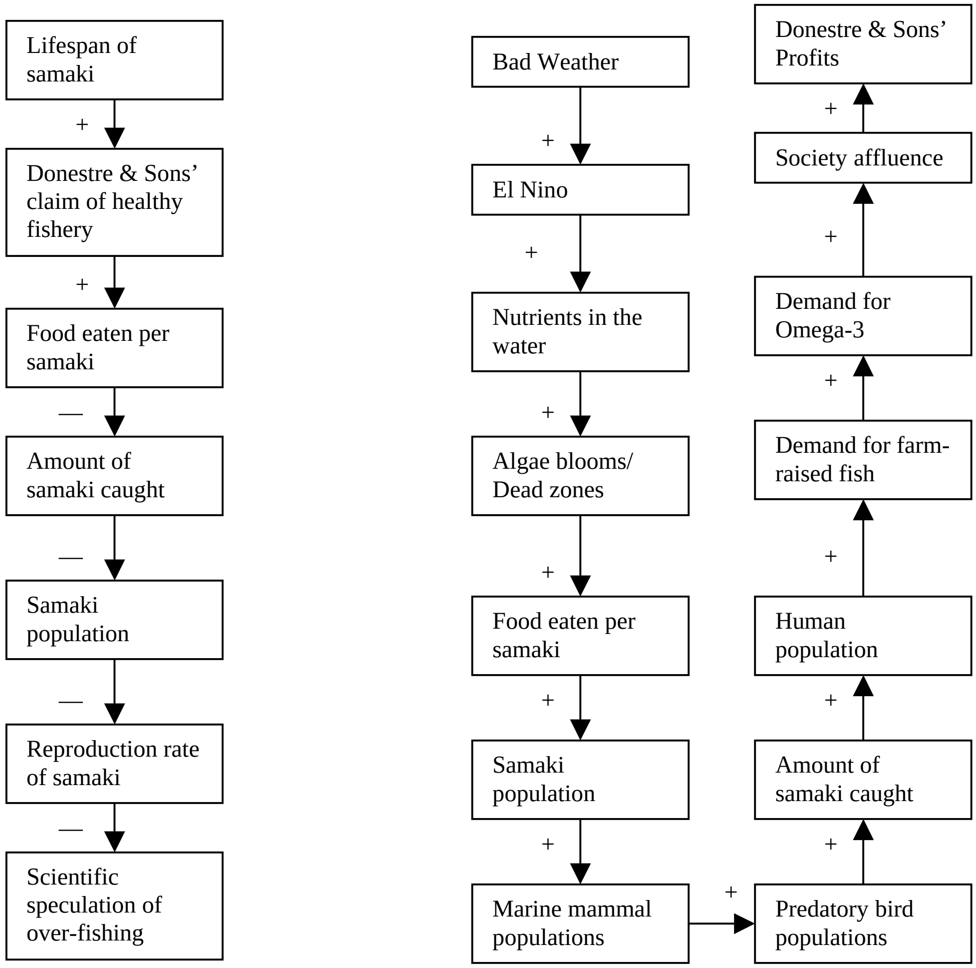


Figure 6-2. Sample causal maps with equivalent link densities



(a) Causal Map from B7 (b) Causal Map from B8

Figure 6-3. Linear causal maps with unlikely connections

CHAPTER 7
EPILOGUE

7.1 Dohaku's Story

There is an old Japanese tale of a samurai, Dohaku, who was returning home from a long trip. Dohaku had nearly reached his own town when he was attacked by two thieves. The attackers cut off Dohaku's head and began rummaging through his bag. But Dohaku was tougher than most. After defeating his assailants, he collected his head and went home to receive medical attention.

When the town's doctor saw Dohaku standing in the doorway with his head tucked under his arm, he knew the situation was grave. Acting quickly, the doctor sewed Dohaku's head back on and placed him in a barrel of rice up to his chin. "Do not move from this barrel," advised the doctor. "I will return to check on you in the morning."

The next morning, the doctor returned and asked Dohaku how he felt. "I feel fine," answered Dohaku. "You are a genius. I would like to leave this barrel. There is much to be done." But the doctor refused, ordering Dohaku to remain in the rice another day. On the second morning, the doctor returned and again Dohaku reported that he felt fine and asked to be set free from the rice. The doctor still refused and ordered another day in the barrel.

And on this went for several days with Dohaku becoming increasingly displeased with his confinement. Then one morning, when the doctor inquired about Dohaku's condition, Dohaku responded, "I feel terrible. I'm hungry. I'm thirsty. My back aches. My knees are torturing me. My whole body feels as if it is on fire."

"Now," said the doctor, "you may leave the barrel."

7.2 Recommendations

Complex systems abound. They affect every aspect of our lives, from our social and economic endeavors down to the food we eat and the air we breathe. This is nothing new. Complex systems have been around far longer than humans. The difference today is the role that humans play in those systems. For the first time in history, the effects of human activity on large-scale environmental systems is occurring rapidly enough to be observed over only a few decades. Peter Vitousek (1994) catalogs these large-scale changes effected by humans, including increase of greenhouse gases, alteration of the global nitrogen cycle, and widespread change of landcover/landuse. He observes, "We're the first generation with the tools to see how the Earth system is changed by human activity; at the same time, we're the last with the opportunity to affect the course of many of those changes" (1994, p. 1873). The "tools" to which Vitousek referes are the scientific instruments used to collect data, and he is quite right to point out our ability to use these tools to document globally significant changes in the Earth system. What we lack are the cognitive tools to recognize our role in these changes and their affect on us.

In this sense, we are akin to Dohaku, who could not feel the damage done to his body because he had lost his connection to it. Similarly, we lack the ability to understand—and at times even to perceive—our connections to the environmental systems that sustain us. Thus, educators teaching with an eye toward sustainability education face a challenge analogous to that of Dohaku's doctor. But the good doctor only had to reconnect a head; educators face the task of reconnecting an entire society. The research here is an attempt to better understand how educators might facilitate an awareness and understanding of those connections.

Systems-oriented instruction represents a promising direction in the context of education for sustainability for two reasons. First, as we have seen in Chapter 2, parties unaffiliated with environmental education have called for the development of curricula that better enable students to deal with the complexity of life in the 21st Century. In this context, systems-oriented education may meet less resistance than previous attempts to include some form of environmental education into curricula. And second, we have proven ourselves—in practice and in numerous academic studies—to be profoundly incapable of comprehending the dynamically complex systems on which we rely everyday.

There are numerous anecdotal accounts of the usefulness of systems-oriented instruction for addressing this problem. However, convincing more educators and administrators to explore systems-oriented instruction will require quantitative data, and any methodology used to acquire that data will likely need to have the following attributes:

1. Short assessment period: Standardized tests are popular, largely because they are easy to administer in large numbers. A number of studies (e.g. Maani and Maharaj 2004) utilize methodologies of assessing systems-oriented instruction that are useful as early explorations, but are too time consuming—for both the researchers and the participants—to be used on a broader scale. Producing the amount of data necessary to draw strong conclusions about the effectiveness of systems-oriented instruction will require efficiency and ease of implementation. It is a truism in education that teachers are always strapped for time. Designing an assessment methodology that requires large periods of class time for assessment is one of the best ways to ensure it is not widely adopted.

2. Broadly applicable design: Existing attempts to implement systems-oriented instruction span a broad range of subjects and educational levels. The methodology, therefore, needs to be usable in a variety of contexts.

3. Focus on learning: Systems thinking is a skill, so any assessment of it should focus on evaluating not what students know, but how they learn.

4. Separation between curriculum design and measurement: Results from the portion of the present study involving the undergraduates were somewhat compromised by the connection between the course objectives and the assessment. The Portland results can be considered more reliable because the students likely did not have any idea how I was going to assess the cognitive maps.

5. Objective evaluation: Some earlier studies assessing systems-oriented instruction rely heavily on feedback from the participants themselves. This feedback is important, particularly in situations where the participants are interested in taking an active role in their own education. However, such data will likely not be as persuasive to teachers and school administrators as objectively calculated cores and indices analogous to those included in standardized tests.

The methodology presented in this report meets these five criteria. Moreover, the results from these studies provide evidence that systems-oriented instruction may live up to its promise as a pedagogical tool for improving students' abilities to comprehend the dynamically complex systems around them. Therefore, I present this methodology as a starting point for the development of a standard assessment of systems thinking programs.

But this is only a starting point. The comparisons made in the present study are rather crude: systems-oriented instruction versus conventional instruction. Long-term research plans should include more nuanced comparisons that explore the advantages and disadvantages of various approaches toward systems-oriented instruction. Such comparisons are difficult to make at present due to the dearth of systems-oriented programs. Thus, it is my hope that the report will add momentum to a positive feedback loop (a virtuous cycle) involving systems-oriented programs and the assessment of those programs like that in Figure 7-1. Repeated results like those reported here would provide persuasive evidence encouraging other educators to adopt system-oriented instruction. And these additional systems-oriented programs would, in turn, provide opportunities for further evaluations that enable us to make more nuanced comparisons.

The samurai Dohaku had a full barrel of rice to keep him safe from harm until he was able to repair the connections that had been severed. We do not have such a luxury. The ever-increasing pressures that we put on our environmental systems are draining our barrel of rice, while our schools do little to provide their students with the skills to understand the connections within complex social and ecological systems. David Orr suggests that in a sustainable society "laws, institutions, and customs would reflect an awareness of interrelatedness, exponential growth, feedback, time delays, surprise, and counterintuitive outcomes" (2002, p. 178). Currently, our educational institutions do not reflect this awareness, and the need for a concerted and organized effort to develop curriculum that fosters understanding of today's complex environmental challenges deserves far more attention than it currently gets.

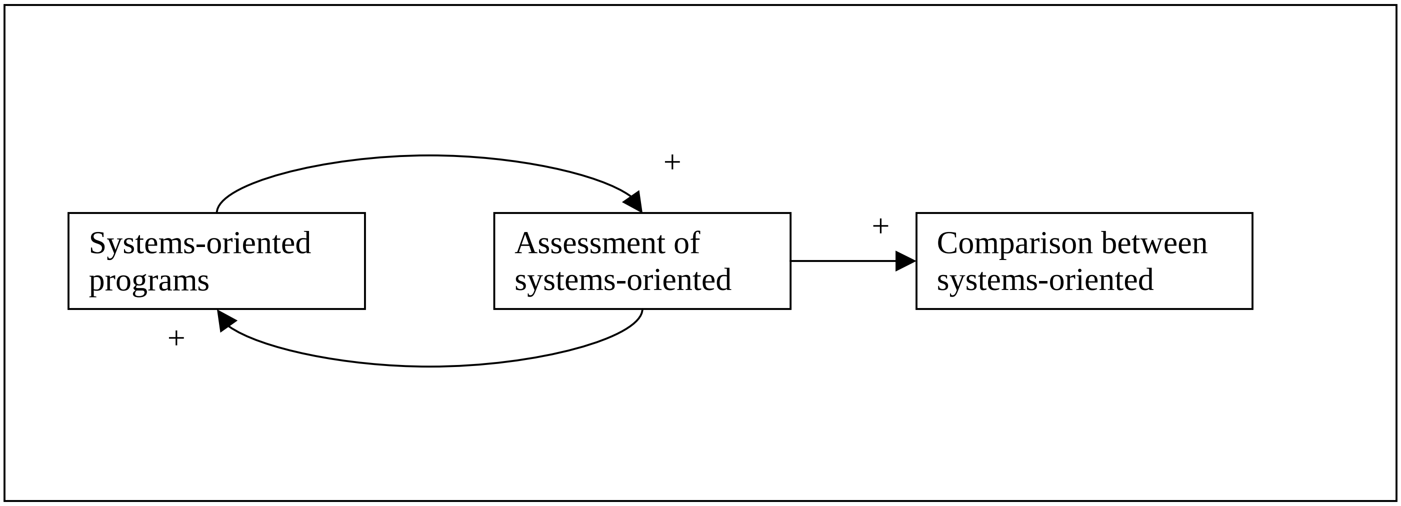


Figure 7-1. Positive Feedback Loop between Systems-oriented Instruction and Assessment

APPENDIX A
ARTICLE USED FOR MAPPING EXERCISE

Samaki:

An Ecological and Industrial Treasure

Samaki, a larger cousin of the herring, may be the most important fish you've never heard of. No one eats samaki directly. They are oily fish filled with tiny bones. However, they still find their way to our dinner tables through indirect routes.

First, they provide an ecological link between microscopic plankton and large predatory fish. In other words, they eat the plankton and, in turn, become an important source of fat-rich, high-nutrient food for many large, high-valued fish, including bass, cod, swordfish, bluefish, and tuna.

Second, they are ground up, dried, and used to make feed for poultry, pigs, beef, and farmed fish, providing an important source of protein for these farm-raised animals as well.

Recently, a third path has been added. The American Heart Association’s declaration in November 2002 that people should consume Omega-3 fatty acids to prevent heart disease has led to an increase in the use of these oils as food supplements for humans as well.

The usefulness of samaki in animal feed and food supplements has made catching samaki big business. Literally millions of pounds of samaki are caught every year for industrial use. In fact, samaki has become so useful to industry that some suggest this once abundant fish is in danger of being overfished.

For example, analyst Carl Kingsolver explains that leaders in the samaki industry have experienced decreased profits and even losses over the past decade. He suggests that this is one sign that the samaki population has been significantly reduced.

However, Donestre & Sons—the largest supplier of samaki-based fish meal and fish oil products in the United States and a major competitor in the world market—suggests that the North American samaki population is doing fine. They explain that the reduced profits were due to factors that were not related to the health of the fishery.

First, prices of their products are highly affected by the amount of fish caught by their international competitors, primarily from South America. High catch levels in South America lower the price of Donestre & Sons’ products.

Second, Donestre & Sons must compete with alternative products. For example, unusually low prices in soybeans and vegetable oil from 1999 to 2001 lowered Donestre & Sons’ sales for those years. Donestre officials point out that the availability of these competing products drove down the price of their own products. The result was a decrease in their profits.

Of course low catches have added to Donestre & Sons’ problems, but company officials explain that these low catches are due to bad weather, rather than to decreases in the samaki population itself. For example, an unusually severe September hurricane season in the Gulf of Mexico in 2000 held back fishing activities and resulted in fewer fish caught.

Lou Robbins, Donestre's Director of Fishing Operations, points out that their international competitors face similar difficulties. In 1992 El Nino events disrupted operations of South American fish suppliers, resulting in higher prices internationally.

Moreover, explains Gregory Murdock, Donestre's Director of Investor Relations, from 2002 until the present, Donestre and Sons has reported net profits every year.

Murdock attributes this recent success to an international increase in aquaculture (fish farming). Soy meal can be used as a substitute for fishmeal in feed for pigs, poultry, and cattle, but farm-raised fish require the protein of fish oils. Therefore, an increase in demand from aquaculture will mean more profits for Donestre & Sons rather than soy meal providers.

A second factor in Donestre's recent success is the increased demand for Omega-3 fatty acid dietary supplements. The American Heart Association announced in November 2002 that people should consume Omega-3 fatty acids to reduce the chance of heart disease. This announcement resulted in a large increase in demand by health-minded consumers.

As a result of these trends, the price of samaki has increased greatly, allowing Donestre & Sons to earn high profits despite lower than expected fish catch for 2003 and 2004.

But despite Donestre & Sons' claims, many people outside the fish-meal industry are concerned about possible decreases in the samaki population, because samaki are also important ecologically. Samaki are an important source of food for larger fish, birds, and marine mammals.

For example, many fishermen suggest that smaller samaki populations have resulted in smaller populations of the highly prized commercial fish that feed on samaki. Even in cases where commercial fish numbers are still high, some people suggest that a lack of samaki has resulted in fish with smaller weight to length ratios. These "skinny" fish are believed to be less hardy due to malnutrition.

Scientists have also linked decreases in samaki populations in Gardner's Bay to declines in the osprey population on Gardner's Island, near the eastern tip of Long Island, New York.

In addition to being an important source of food, samaki are filter feeders that help control algal growth in coastal waters. Adults can be seen swimming in tightly packed groups with their mouths open, filtering out algae species that many other fish won't eat.

Some people place great importance in the samakis' feeding habits. For example, declines in samaki populations have been linked to algal blooms along coastal waters. That's where the algae grows so thick that it kills everything else in the area, producing what's called "dead zones." In this sense, samaki not only provide a food source for valuable fish species, but also make the coastal waters a better place for fish to live.

Because of the ecological importance of samaki, some argue that the federal government should pass laws that manage or limit the amount of samaki that can be caught. Others call for incorporating limits to samaki fishing into more strict management of the coastal ecosystems as a whole. Companies like Donestre & Sons have said they would oppose such laws.

Fishermen will tell you that their industry offers very little certainty. But perhaps one thing is definite: the health and management of the samaki fishery will be hotly debated over the coming years.

APPENDIX B
SCRIPT USED FOR MAPPING EXERCISE

Part 1:

Thank you for participating in this study on how people interpret new information. You have an article in your packet. In a moment, I will give you ten minutes to read the article. After that time is up, you will put that article aside, and we will go through a short series of exercises relating to the article. Since you will not have the article in front of you, you may want to jot down a few notes. That's fine, but you do not have to. I am more interested in you general impression of the article, than I am in specific details. I can help you to remember any specific names you might need.

Are there any questions?

Begin reading. You have ten minutes. (Wait eight minutes.)

You have two minutes left. (Wait two minutes.)

Now, imagine you are presenting the issue of samaki to a class in which the students know nothing about the issue. What aspects of the situation would you relate to the students so that they would get the most complete understanding of the issue? In your packet, you have a clipped stack of forty cards. Please take it out. Go through the stack of cards and pick out any cards that express concepts or things you would discuss with the class. The goal here is to give the class the fullest understanding possible. But, choose only those cards that express things you can explain—that is, things you understand. You may pick as many cards as you like. You will notice that there are five blank cards at the bottom of you stack. Keep these handy. You will use them soon.

Once you have gone through the stack and chosen all the cards that you want, put the paper clip back on the cards that were NOT chosen. This will keep them out of the way. Keep the blank cards out for now.

Spread the cards you chose out in front of you. Looking at the cards, do you feel that anything has been missed? If so, use the blank cards to include any aspects of the issue that are not yet covered. You do not have to use the blank cards. Place any unused blank cards back into the clipped stack.

[Wait for students to finish separating cards.]

Now I'd like you to place your chosen cards into groups. You may use any criteria you like to form your groups, and a group may include any number of cards.

[Wait for students to finish grouping cards.]

Now, I would like you to give each group a short, descriptive title. Take the "Response Form" from your packet. Look at your first group of cards and think of a label for that group. Write the label on the Response Form in the space marked "Group A." Then write an "A" on each of the cards in that group in the space provided on the upper right corner of each card.

Repeat these steps for the rest of the group.

[Wait for students to label groups.]

Take one last look at your groups. If you feel comfortable with the way you have grouped the cards, then you have finished the first part of the exercise.

Part 2

For the second part of the exercise, we will look at cause-and-effect relationships between cards. To do this, you will need the large piece of paper in your packet and a glue stick. On this sheet of paper you will make a "map" of the situation you just read about.

This is a specific kind of map, so I'll have to give you a quick lesson on how to make one. It's easiest to learn with an example, so let's take the example of a popular television show. Someone name a popular show. [Note name of show.]

On the day after the show do people usually talk about it at school?

Now, we'll map that situation. Say I have two cards. Card 1 is "Number of People Who Watch the Show" and Card 2 is "Number of People Who Talk about the Show at School."

Number of People Who Watch ______________	Number of People Who Talk about _________ Later

Looking at these two cards, tell me which of the following sentences best expresses their relationship.

[Written on chalk board]

An increase in Card 1 causes an increase in Card 2. [Positive Arrow]

An increase in Card 1 causes a decrease in Card 2. [Negative Arrow]

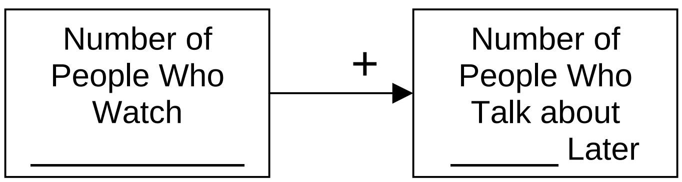
An increase in Card 1 does not directly affect Card 2.

If I chose Sentence one, then I will connect the two cards with a positive arrow going from Card 1 to Card 2. (Show Arrow)

If I chose Sentence 2, then I will connect the two cards with a negative arrow going from Card 1 to Card 2. (Show Arrow)

If I chose Sentence 3, then I will not connect the two cards.

Since Sentence 1 describes the situation, I will connect the two cards with a positive arrow. Like this.

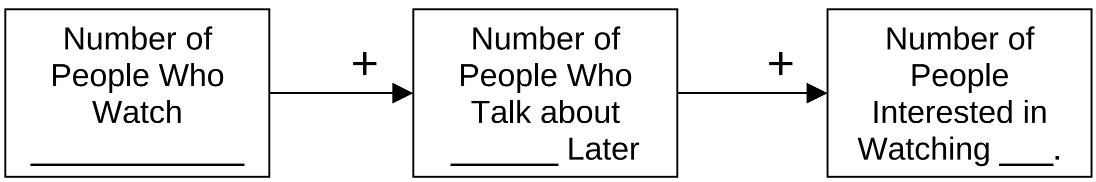


Note that the positive arrow does not always mean an increase. It just means that the two cards move in the same direction. So, according to this picture, a DECREASE in the “Number of People Who Watch ______” causes a DECREASE in the “Number of People Who Talk About the Show” the next day. The main thing is that the two cards move in the same direction.

Now, if you’re friends are all talking about how much they like a television show that you have never seen, would you probably be at least a little bit interested in watching the show too?

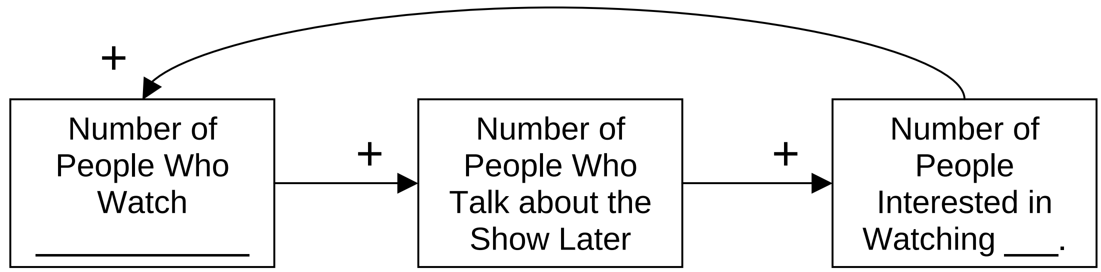
Then let's add another card: "Number of People Interested in Watching _____." I might suggest that Sentence 1 would also describe the relationship between the "Number of People Who Talk about the Show" and the "Number of People Interested in Watching _____." In other words, the more people who are talking about the show, the more other people become interested in watching the show. I would show that by drawing another positive arrow, like this: :

[Written on board]

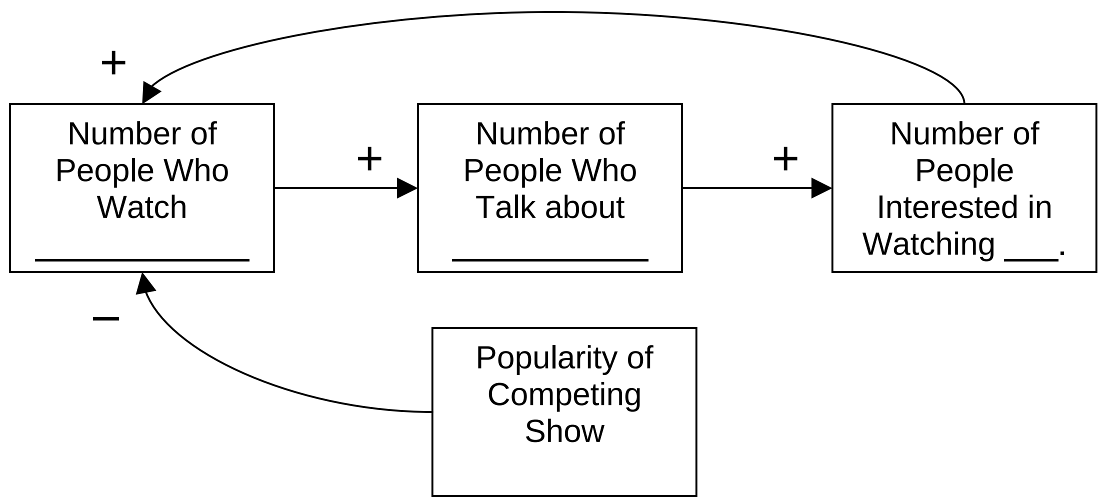


Note that according to this map, an increase in the "Number of People Who Watch ___" will eventually lead to an increase in the "Number of People Interested in Watching ___." (Show path from Card 1 through Card 2 to Card 3.) However, since the "Number of People Who Watch ___" does not affect the "Number of People Interested in Watching ___" directly, no arrow is necessary from Card 1 straight to Card 3.

But arrows can point in any direction. For example, if I want to say that the more "People Interested in Watching," the more "People will watch," then I could do so with a positive arrow from Card 3 to Card 1, like this:



There's one last thing I want to show you before you start making your own maps: negative arrows. Suppose another show that airs at the same time becomes popular. That show's popularity would likely decrease the number of people who watch __________:



This implies that the more popular the competing television show becomes, the fewer people there will be watching _____________.

This has just been a small example. As you make your own map, there is no limit to the number of connections you can include.

Now, use the sheet of sketching paper in front of you and the cards you've chosen to make a map of the samaki issue. I recommend not gluing the cards down right away because you may want to move them around as you make your map.

Start this part by picking any two cards at random—Card 1 and Card 2. Looking at the cards, choose one the following sentences that best expresses the relationship between the two cards:

An increase in Card 1 causes an increase in Card 2 (Positive arrow)

An increase in Card 1 causes a decrease in Card 2. (Negative arrow)

An increase in Card 1 does not directly affect Card 2. (No arrow)

[Show the sentences on a blackboard with the arrows drawn.]

Here are a few quidelines to follow.

All of your chosen cards must be included on the map. If you do not know how to connect it to other cards, then you need to question whether you should keep the card.

If, when making your map, you find that you need to include cards you discarded earlier, you may do so. However, please show which group—from the groups you created earlier—this card would go into.

Be sure to label each arrow positive of negative, depending on the relationship between the two cards.

Be careful to show all direct connections that you see. Try not to forget a connection simply because two cards are far apart on your sheet of paper.

Please take care to make sure each connection is clear. (On particularly busy maps, you may find that dotted lines help to distinguish one connection from another.)

Once you have completed the map, take a moment to review the connections. Does this map accurately illustrate your view of the issue? If so, you are done with the mapping portion of this exercise.

For the final portion of the exercise, please answer the following set of questions:

Do you feel that samaki are in danger of being overexploited? Please explain your answer. If you do not feel informed enough to have an opinion, what else would you want to know about this issue in order to form your own opinion about what should be done with the samaki?

Thank you for your participation in this study.

LIST OF REFERENCES

Ackoff RL. 1999. *Ackoff's Best: His Classic Writings on Management*. John Wiley & Sons: New York.

Addams J. 1902. *Democracy and Social Ethics*. MacMillan: New York.

Agyris C. 1976. *Increasing Leadership Effectiveness*. John Wiley & Sons: New York.

Akimoto H. 2003. Global air quality and pollution. *Science* **302**(5649): 1716-1719.

Allesina S. 2000. *Windows Application for Network Analysis Digraphs*. Computer Software. http://www.dsa.unipr.it/~alle/ena. August 2006.

Ambrosini V, Bowman C. 2001. Tacit knowledge: Some suggestions for operationalization. *Journal of Mangement Studies* **38**(6): 811-829.

Amtmann L. 1996. *Stakeholder perspectives on wild and scenic rivers: a new cognitive tool for enhancing participation*. M.S. Thesis. The University of Michigan. Ann Arbor, MI.

Axelrod R. 1976. *Structure of Decision: Cognitive Maps of Political Elites*. Princeton University Press: Princeton, NJ.

Banathy B. 1991. *Systems Design of Education: A Journey to Create the Future*. Educational Technology Publications: Englewood Cliffs, NJ.

Barnett SM, Ceci SJ. 2002. When and where do we apply what we learn?: a taxonomy for far transfer. *Psychological Bulletin* **128**(4): 612-637.

Bateson G. 1979. *Mind and Nature: A Necessary Unity*. Bantam Books: New York.

--. 1987. *Steps to an Ecology of Mind*. Jason Aronson, Inc.: New Jersey.

Bell BI. 1949. *Crisis in Education: A Challenge to American Complacency*. Whittlesey House: New York.

Bellack A, June A. 1969. History of curriculum thought and practice. *Review of Educational Research* **39**(3): 283-292.

Benjamin HRW. 1939. *Saber-Tooth Curriculum, including Other Lectures in the History of Paleolithic Education*. McGraw-Hill: New York.

Berliner DC, Bingman VP. 1997. *The Manufactured Crisis: Myths, Fraud, and the Attack on America's Public Schools*. Longman Publishers: White Plains, NY.

Bersier L-F, Sugihara G. 1997. Scaling regions for food web properties. *Proceedings of the National Academy of Sciences* **94**: 1247-1251.

Bertalanffy Lv. 1949. *Problems of Life*. John Wiley & Sons, Inc.: New York.

--. 1969. *General Systems Theory: Foundations, Development, Applications*. George Braziller: New York.

Berverly T, Gnyawali D. 2002. Mapping managers' market orientations regarding new product success. *Journal of Product Innovation Management* **19**(4): 259-276.

Bestor A. 1953. *Educational Wastelands: A Retreat From Learning in Our Public Schools*. University of Illinois Press: Urbana, IL.

Bingman VP, Siegel JJ, Gagliardo A, Erichsen JT. 2006. Representing the richness of avian spatial cognition: Properties of a lateralized homing pigeon hippocampus. *Reviews in the Neurosciences* **17**(1-2): 17-28.

Bodin C, Benhamou S, Poulle M-L. 2006. What do European badgers (Meles meles) know about spatial organization of neighboring groups. *Behavioral Processes* **72**(1): 94-99.

Booth Sweeney L, Meadows D. 2001. *Systems Thinking Playbook*. Institute for Policy and Social Science: Durham, NH.

Booth Sweeney L, Sterman J. 2000. Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review* **16**(4): 249-286.

Bougon M, Baird N, Komocar JM, Ross W. 1983. Identifying strategic loops: The self-Q interviews In *Mapping Strategic Thought*. Huff, A., Ed. Wiley & Sons, Ltd.: Chichester, West Sussex, England.

Bougon M. 1992. Congregate cognitive maps: A unified dynamic theory of organization and strategy. *Journal of Management Science* **29**(3): 369-389.

Boulding K. 1953. Toward a general theory of growth. *Canadian Journal of Economics and Political Science* **19**(3): 326-340.

Bowers CA. 1995. *Educating for an Ecologically Sustainable Culture*. SUNY Press: New York.

Braun W. 2002. *The System Archetypes*. www.uni-klu.ac.at/~gossimit/pap/sd/wb_sysarch.pdf. July 2006.

Brehmer B. 1980. In one word: Not from experience. *Acta Psychologica* **45**: 223-241.

--. 1992. Decision-making: Human control of complex systems. *Acta Psychologica* **81**: 211-241.

Broadbent D. 1977. Levels, hierarchies, and the locus of control. *Quarterly Journal of Experimental Psychology* **29**: 181-201.

Brown AD. 2003. *Feed or Feedback: Agricultural Population Dynamics and the State of the Planet*. International Books: Utrecht, Netherlands.

Bruner J. 1953. *The Process of Education*. Harvard University Press: Cambridge, MA.

Bush GW. 2006. *State of the Union Address*. http://www.whitehouse.gov/stateoftheunion/2006/index.html. United States Capital: Washington, DC. July 2006.

Capra F. 1982. *The Turning Point: Science, Society, and the Rising Culture*. Bantam Books: New York.

Cavaleri S, Sterman J. 1997. Towards evaluation of systems thinking interventions: A case study. *System Dynamics Review* **13**(2): 171-186.

Clark I, Mackaness W. 2001. Management 'intuition': an interpretive account of structure and content of decision schemas using cognitive maps. *Journal of Management Studies* **38**(2): 147-172.

Cohen J, Bussey K. 2003. Rats form cognitive maps from spatial configurations of proximal arm cues in an enclosed 4-arm radical maze. *Learning and Motivation* **34**(2): 168-184.

Cohen J. 2003. Human population: The next half century. *Science* **302**(5648): 1172-1175.

Cohen LJ. 1981. Can human irrationality be experimentally demonstrated. *The Behavioral and Brain Sciences* **4**: 317-331.

Colorado Department of Education. 1995. *Colorado Model Content Standards for Science*. http://www.cde.state.co.us/cdeassess/documents/standards/sci.htm#introduction. April 2006.

Council of Economic Advisors. 2005. *The Economic Report of the President*. Goverment Printing Offices: Washington, DC.

Cremin L. 1955. The revolution in American secondary education. *Teachers College Record* **56**(6): 295-308.

--. 1961. *The Transformation of Hte School: Progressivism in American Education*. Alfred A. Knopf: New York.

--. 1971. Curriculum-making in the United States. *Teachers College Record* **73**(2): 207-220.

Daly H. 1973. *Towards a Steady-State Economy*. Freeman: San Francisco.

Davis O.L. 1976. *Perspectives on Curricular Development 1776-1976*. Association for Supervision and Curriculum Development: Washington, DC.

Dewey J. 1902. *The Child and the Curriculum*. University of Chicago Press: Chicago.

--. 1916. *Democracy and Education: An Introduction to the Philosophy of Education*. MacMillan: New York.

Doyle JK. 1997. The cognitive psychology of systems thinking. *System Dynamics Review* **13**: 253-265.

Doyle JK, Radzicki M, and Trees WS. 1998. *Measuring change in mental models of dynamic systems: An exploratory study*. http://www.wpi.edu/Academics/Depts/SSPS/Research/Papers/14.pdf#search=%22Measuring%20Change%20in%20Mental%20Models%20of%20Dynamic%20Systems%3A%20An%20Exploratory%20Study%22. August 2006.

Dumaine B. 1989. What the leaders of tomorrow see. *Fortune* **120**(1): 48-54.

Dunne J, Williams R, Martinez N. 2002. Network structure and biodiversity loss in food webs: Robustness increases with connectance. *Ecology Letters* **5**: 558-567.

Eden C. 1992. On the nature of cognitive maps. *Journal of Management Studies* **29**: 261-265.

Eden C, Ackerman F, Cooper S. 1992. The analysis of cause maps. *Journal of Management Studies* **29**: 309-324.

Eden C, Ackerman F. 1998. *Making Strategy: The Journey of Strategic Management Thought*. Sage: London.

Edwards W. 1954. Theory of decision-making. *Psychological Bulletin* **51**: 380-417.

--. 1961. Behavioral decision theory. *Annual Review of Psychology* **12**: 473-498.

--. 1962. Dynamic decision theory and probabalistic information processing. *Human Factors* **4**: 59-73.

Einborn HJ, Hogarth RM. 1978. Confidence and judgment: Persistence of the illusion of validity. *Psychological Review* **85**(5): 395-416.

Eliot C. 1892. Wherein popular education has failed. In *Charles W. Eliot and Popular Education*. Krug, Edward A., Ed. Teachers College, Columbia University: New York.

Ellard CG. 2000. Landmark navigation in gerbils: The role of the posterior cortex. *Pscyobiology* **28**(3): 325-338.

Ellis AK. 2004. *Exemplars of Curriculum Theory*. Eye on Education, Inc.: Larchmont, NY.

Essingon TE, Beaudreau A, Wiedenmann J. 2006. Fishing through marine food webs. *Proceedings of the National Academy of Sciences of the United States of America* **103**(9): 3171-3191.

Evans J. 2006. *Kings Bay/Crystal River Florida: A conceptual proposal to restore a submerged macrophyte ecosstem through adaptive management*. Dissertation. University of Florida. Gainesville, FL.

Everard M. 1999. Towards sustainable development of still water resources. *Hydrobiologia* **395**: 29-38.

Ezzati M, Utzinger J, Cairncross S, Cohen AJ, Singer BH. 2005. Environmental risks in the developing world: Exposure indicators for evaluating interventions, programmes, and policies. *Journal of Epidemiology and Community Health*. 59: 15-22.

Fath BD, Patten BC. 1999. Review of the foundations of network environ analysis. *Ecosystems* **2**: 167-179.

Fiol CM, Huff AS. 1992. Maps for managers: Where are we? Where do we go from here? *Journal of Management Studies* **29**: 267-285.

Florida Department of Education. 2005. *Sunshine State Standards*. http://www.firn.edu/doe/curric/prek12/frame2.htm. Tallahassee, FL. August 2006.

Ford J, Hegarty WH. 1984. Decision makers' belief about the causes and effects of structure: An exploratory study. *Academy of Management Journal* **27**(2): 271-291.

Forrester JW. 1961. *Industrial Dynamics*. MIT Press: Cambridge, MA.

--. 1969. *Urban Dynamics*. MIT Press: Cambridge, MA.

--. 1992. *System Dynamics and Learner-Centered Learning in Kindergarten through 12th Grade Education*. http://sysdyn.clexchange.org/sdep/Roadmaps/RM1/D-4337.pdf. Creative Learning Exchange: Cambridge, MA. May 2006.

--. 1996. *System Dynamics and K-12 Teachers*. Lecture at the University of Virginia. May 5.

Fry S, Rudiger W. 2002. Honey bees store landmarks in an egocentric frame of reference. *Journal of Comparative Physiology* **187**(12): 1009-1016.

Funke J. 1991. Solving complex problems: Exploration and control of complex systems In *Complex Problem Solving: Principles and Mechanisms*. Sterberg, R. J. and Frensch, P. A., Ed. Erlbaum Associates: Hillsdale, NJ.

G8 Gleneagles. 2005. *Climate change, clean energy and sustainable development*. http://www.g8.gov.uk. July 2006.

Gentner D, Gentner DR. 1983. Flowing waters or teeming crowds: Mental models of electricity In *Mental Models*. Gentner, D. and Stevens, A. L., Ed. Earlbaum Associates: Hillsdale, NJ.

Gigerenzer G. 1991. How to make cognitive illusions disappear: Beyond 'heuristics and biases' In *European Review of Social Psychology, Vol. 2*. Stroebe, Wolfgang and Hewstone, Miles, Ed. Wiley Publishing: West Sussex.

--. 1996. On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review* **103**(3): 592-596.

Gigerenzer G, Todd PM. 1999. *Simple Heuristics That Make Us Smart*New York.

Gleik PH. 2003. Global freshwater resources: Soft-path solutions for the 21st century. *Science* **302**(5650): 1524-1528.

GLOBE. 2005. *The Next Generation Globe*. GLOBE Program: Boulder, CO.

Gore, A. 2006. Interview with Terry Gross. *Fresh Air*. 30 May. National Public Radio.

Gunderson L, Holling CS. 2001. *Panarchy: Understanding Transformations in Systems of Humans and Nature*. Island Press: Washington, DC.

Hardin G. 1968. The tragedy of the commons. *Science* **162**(3859): 1243-1248.

Hargreaves A, Goodman I. 1999. Preface In *Changing Leadership for Changing Times*. Leithwood, K, Ed. Open University Press: Philadelphia.

Hoch SJ, Deighton J. 1989. Managing what consumers learn from experience. *Journal of Marketing* **53**: 1-20.

Hodgkinson G, Maule AJ, Brown N. 2004. Causal cognitive mapping in the organizational strategy field: A comparison of alternative elicitation procedures. *Organizational Research Methods* **7**(1): 3-26.

Hofstadter R. 1963. *Anti-Intellectualism in American Life*. Knopf: New York.

Hogarth RM. 1980. *Judgement and Choice: The Psychology of Decision*. Wiley Publishing: New York.

Holmgren D. 2004. *Essence of permaculture*. http://www.holmgren.com.au/DLFiles/PDFs/EssenceofPC3.pdf. Holmgren Design Services: Hepburn, Victoria. July 2006.

Huff AS. 1990. Mapping strategic thought In *Mapping Strategic Thought*. Huff, A. S., Ed. John Wiley & Sons: New York.

Huz S, Andersen DF, Richardson GP, Boothroyd R. 1997. A framework for evaluating systems thinking interventions: An experimental approach to mental health system change. *System Dynamics Review* **13**(2): 149-169.

Jackson P. 1992. Conceptions of curriculum and curriculum specialists In *Handbook of Research on Curriculum*. Jackson, Phillip, Ed. MacMillan: New York.

Jacobs LF. 2003. The evolution of the cognitive map. *Brain Behavior and Evolution* **62**(2): 128-139.

Jenkins M, Johnson G. 1997. Linking managerial cognition and organizational performance: A preliminary investigation using causal maps. *British Journal of Management* **8**(Special Issue): S77-S90.

Jenkins M. 2003. Prospects for biodiversity. *Science* **302**(5648): 1175-1177.

Johnson J, Borgatt S, Luczkovich J, Everett M. 2001. Network role analysis in the study of food webs: An application of regular role coloration. *Journal of Social Structure* **2**(3): 1-35.

Jung C. 1959. Archetypes of the collective unconscious In *The Basic Writings of C.G. Jung*. de Laszlo, Violet S., Ed. The Modern Library: New York.

Kahneman D, Tversky A. 1974. Judgment under uncertainty. *Science* **185**(4157): 1124-1131.

--. 1981. The framing of decisions and the psychology of choice. *Science* **211**: 453-458.

--. 1996. On the reality of cognitive illusions. *Psychological Review* **109**(3): 582-591.

Kahneman D, Slovic P, Tversky A. 1982. *Judgment Under Uncertainty*. Cambridge University Press: New York.

Kainz D, Ossimitz G. 2002. Can students learn stock–flow thinking? An empirical investigation. *Proceedings of the 2002 International System Dynamics Conference*. System Dynamics Society: Albany, NY.

Kardaras D, Karakostas B. 1999. The use of fuzzy cognitive maps to simulate information systems strategic planning process. *Information and Software Technology* **41**(4): 197-210.

Kearney A, de Young R. 1995. A knowledge-based internvention of promoting carpooling. *Environment and Behavior* **27**(5): 650-678.

Kearney A, Kaplan S. 1997. Toward a methodology for the measurement of knowledge structures of ordinary people: the conceptual content causal map (3CM). *Environment and Behavior* **29**(5): 579-617.

Kearney A. 1999. *Teacher perspectives on environmental education and school improvement: A report for the Evergreen Center for Education and Improvement*. Evergreen State College: Olympia, WA.

Kearney A, Bradley G, Kaplan R, Kaplan S. 1999. Stakeholder perspectives on appropriate forest management in the Pacific Northwest. *Forest Science* **45**(1): 62-73.

Kempton W. 1986. Two theories of home heat control. *Cogntive Science* **10**(1): 75-90.

Kleibard H. 2005. *The Struggle for the American Curriculum*, Third Edition. Routledge: New York.

Knapp D. 2000. The Thessalonki Declaration: A wake-up call for environmental education. *Journal of Environmental Education* **31**(3): 32-39.

Koffka K. 1935. *Principles of Gestalt Psychology*. Harcourt Brace: New York.

Kuhn T. 1987. What are scientific revolutions In *The Probabalistic Revolution*. Kruger, L, Daston, L., and Heidelberger, M., Ed. Cambridge University Press: Cambridge.

Kuhn T. 1996. *The Structure of Scientific Revolutions,* Third Edition Edition. University of Chicago Press: Chicago.

Langfield-Smith K. 1992. Exploring the need for a shared cognitive map. *Journal of Management Studies* **29**(3): 349-367.

Larkin J. 1983. The role of problem representation in physics In *Mental Models*. Genter, Dedre. and Stevens, Albert L., Ed. Erlbaum Associates: Hillsdale, NJ.

Laszlo E, Masulli I, Artigiani R, Csanyi V. 1993. *The Evolution of Cognitive Maps: New Paradigms for the Twenty-First Century*. Gordon and Breach Publishers: Amsterdam.

Llinas, Mimo, Conde. 2003. Spatial use and cognitive maps in a group of Guinean baboons (Papio papio) in semi-free conditions. Presented at the *Fifth Meeting of the Spanish Primatological Society*, September 16-20.

Lopes LL. 1991. The rhetoric of irrationality. *Theory and Psychology* **1**: 65-82.

Lotka AJ. 1945. The law of evolution as a maximal principle. *Human Biology* **17**: 167-194.

Lyneis J, Lyneis D. 2003. Bathtub dynamics at WPI. *Proceedings of the 21st International Conference of the System Dynamics Socieity,* New York. System Dynamics Society: Albany, NY.

Maani K, Maharaj V. 2004. Links between systems thinking and complex decision making. *System Dynamics Review* **20**(1): 21-48.

Marshall JD. 2000. *Turning Points in Curriculum: A Contemporary American Memoir*. Merrill: Upper Saddle River, NJ.

Maturana H, Varella F. 1987. *The Tree of Knowledge: The Biological Roots of Human Understanding*. New Science Library: Boston.

McCarthy C, Crichlow W. 2005. *Race, Identity, and Representation in Education,* Second Edition. Routledge: New York.

McCormick J. 1987. The wisdom of Solomon. *Newsweek* **110**(7): 62-63.

Meadows D, Meadows D, Randers J, Behrens WW. 1972. *Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*New York.

Meadows D. 1991. *The Global Citizen*. Island Press: Washington, DC.

Meadows D, Meadows D, Randers J. 1992. *Beyond the Limits: Confronting Global Collapse, Envisioning a Sustainable Future*. Chelsea Green: White River Junction, VT.

Medin D, Bazerman MH. 1999. Broadening behavioral decision research: Multiple levels of cognitive processing. *Psychonomic Bulletin and Review* **6**(4): 533-546.

Meehl PE. 1986. Causes and effects of my disturbing little book. *Journal of Personality Assessment* **50**(3): 370-375.

Meyer A. 1967. *An Educational History of the American People*. McGraw-Hill Book Company: New York.

Moeser SD. 1988. Cognitive mapping in a complex building. *Environment and Behavior* **20**(1): 21-49.

Moore DW, Monaghan EJ, Hartman DK. 1997. Values of literacy history. *Reading Research Quarterly* **32**(1): 90-102.

Moore SB, Ausley LW. 2004. Systems thinking and green chemistry in the textile industry: Concepts, technologies and benefits. *Journal of Cleaner Production* **12**(6): 585-601.

Moxnes E. 2000. Not only the tragedy of the commons: Misperceptions of feedback and policies for sustainable development. *System Dynamics Review* **16**(4): 325-348.

National Commission on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform*. http://www.ed.gov/pubs/NatAtRisk/index.html. United States Government Printing Office: Washington, DC. April 2006.

National Education Association. 1918. The cardinal principles of secondary education In *The American Curriculum: A Documentary History*. Willis, George, Schubert, William H., Bullough, Robert V., Kridel, Craig, and Holton, John T., Ed. Greenwood Press: Westport, CT. 153-162.

National Science Foundation (NSF). 2001. *National Science Foundation Strategic Plan: FY 2001-2006*. http://www.nsf.gov/pubs/2001/nsf0104/start.htm. April 2006.

New York State Education Department. 1999. *Resource Guide with Core Curriculum*. http://www.emsc.nysed.gov/ciai/cores.htm#ss. April 2006.

Nisbett RE, Krantz DH, Kunda Z. 1983. The use of statistical reasoning heuristics in everyday deductive reasoning. *Psychological Review* **90**: 339-363.

Norman DA. 1983. Some observations of mental models In *Mental Models*. Gentner, Dedre and Stevens, Albert L., Ed. Earlbaum Associates: Hillsdale, NJ.

Novak J, Gowin DB. 1984. *Learning How to Learn*. Cambridge University Press: Cambridge.

Novak J. 1998. *Learning, Creating, and Using Knowledge*. Lawrence Erlbaum Associates: London.

Odum HT. 1983. *Systems Ecology: An Introduction* . John Wiley & Sons: New York.

Odum HT. 1994. *Ecological and General Systems*. University Press of Colorado: Niwot, Colorado.

Odum HT, Odum EC. 2001. *A Prosperous Way Down*. University Press of Colorado: Boulder, CO.

Ornstein AC, Levine DU. 2000. *Foundations of Education,* Seventh Edition. Houghton Mifflin Company: Boston.

Orr DW. 1994. *Earth in Mind: On Education, Environment, and the Human Prospect*. Island Press: Washington, DC.

--. 2002. *The Nature of Design: Ecology, Culture, and Human Intention*. Oxford University Press: New York.

Pala O, Vennix JAM. 2005. Effect of system dynamics education on systems thinking inventory task performance. *System Dynamics Review* **21**(2): 147-172.

Pauly D, Alder J, Bennett E, Christensen V, Tyedmers P, Watson R. 2003. The future of fisheries. *Science* **302**(5649): 1359-1361.

Payne JW, Bettman JR, Johnson EJ. 1992. Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology* **43**: 87-131.

Pentland L, Anderson V, Dye S, Wood S. 2003. The nine-box maze test: A measure of spatial memory development in children. *Brain and Cognition* **52**(2): 144-154.

Pinar W, Reynolds W, Slattery P, Taubman P. 1995. *Understanding Curriculum: An Introduction to the Study of Historical and Contemporary Curriculum Discourses*. Peter Lang: New York.

Pinar W. 2003. *Curriculum: Toward New Identities*. Garland Publishers: New York.

Polanyi M. 1969. *Knowing and Being*. University of Chicago Press: Chicago.

Portugali JEd. 1996. *The Construction of Cognitive Maps*. Kluwer Academic Publishers: Dordrecht, MA.

Richmond B. 2000. *The "Thinking" in Systems Thinking: Seven Essential Skills*. Pegasus Communications: Waltham, MA.

Roberts FS. 1976. The questionnaire method In *The Structure of Decision: Cognitive Maps of Political Elites*. Axelrod, Robert, Ed. Princeton University Press: Princeton, NJ.

Rockefeller Brothers Fund. 1958. *The pursuit of excellence: Education and the future of America*. Doubleday & Company: Garden City, NY.

Roos L, Jr., Hall R. 1980. Influence diagrams and organizational power. *Administrative Science Quarterly* **25**(1): 57-71.

Rorty R. 1999. *Philosophy and Social Hope*. Penguin Books: New York.

Scobey M-M, Graham G. 1970. *To Nurture Humaneness*. National Education Association: Washington, DC.

Senge P. 1990. *The Fifth Discipline*. Doubleday Currency: New York.

Senge P, Kleiner A, Roberts C, Ross R, Roth G, Smith B. 1999. *The Dance of Change*. Doubleday Currency: New York.

Shewchuk R, Rivera P, Elliott T, Adams A. 2004. Using cognitive mapping to understand problems experienced by family caregivers of persons with severe disabilities. *Journal of Clinical Psychology in Medical Settings* **11**(3): 141-150.

Silberman C. 1970. *Crisis in the Classroom: The Remaking of American Education*. Random House: New York.

Simon, H. 1956. Rational choice and the structure of the environment. *Psychological Review* 63(2): 129-138.

Smith BO, Stanley WO, Shores JH. 1957. *Fundamentals of Curriculum*, Revised Edition. World Book Company: Yonkers-on-Hudson, NY.

Smith KR, Ezzati J. 2005. How environmental health risks change with development: The epidemiologic and environmental risk transitions revisited. *Annual Review of Environment and Resources* **30**: 291-333.

Spencer H. 1861. *Education: Intellectual, Moral, and Physical*. D. Appleton: New York.

Sterman J. 1989. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision-making experiment. *Mangement Science* **35**(3): 321-339.

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

Sterman J, Booth Sweeney L. 2002. Cloudy skies: Assessing public understanding of global warming. *System Dynamics Review* **18**(2): 207-240.

Stocking MA. 2003. Tropical soils and feed security: The next 50 years. *Science* **302**(5649): 1356-1359.

Strong M. 2001. Forward In *Human Ecology*. Marten, Gerald, Ed. Earthscan: Sterling, VA.

Taba H. 1962. *Curriculum Development: Theory and Practice*. Harcourt, Brace & World: New York.

Tainter J. 1988. *The Collapse of Complex Societies*. Cambridge University Press: Cambridge.

Tanner D, Tanner L. 1980. *Curriculum Development: Theory into Practice*. Merrill: Columbus, OH.

--. 1990. *History of the School Curriculum*. MacMillan: New York.

Thompson L. 2001. *The Mind and Hearth of the Negotiator,* 2nd Edition. Prentice Hall: Upper Saddle River, NJ.

Thorndike EL. 1901. The influence of improvement in one mental function upon the efficiency of other functions: II. the estimation of magnitudes. *Psychological Review* **8**: 384-395.

Toda M. 1962. The design of a fungus-eater. *Behavioral Science* **7**: 164-183.

Tolman EC. 1948. Cognitive maps in rats and men. *Psychological Review* **55**(4): 189-208.

Tully S. 2000. Water, water everywhere. *Fortune* **141**(10): 342-354.

Tyler, BB. Gnyawali DR. 2002 Mapping managers' market orientations regarding new product success, *Journal of Product Innovation Management*19(4): 259-276.

Ulanowicz RE. 2004. Quantitative methods for ecological network analysis. *Computational Biology and Chemistry* **28**: 321-339.

United Nations. 2002. *No Sustainable Development without Education*. http://portal.unesco.org/en/ev.php-URL_ID=5751&URL_DO=DO_TOPIC&URL_SECTION=201.html. Unesco Press: Paris.

United Nations. 2003. *Rsolution 57/254: United Nations Decade of Education for Sustainable Development*. http://daccessdds.un.org/doc/UNDOC/GEN/N02/556/12/PDF/N0255612.pdf?OpenElement .

United States Department of Commerce USDoEUSDoLatNIfLaSBA. 1999. *21st century skills for 21st century jobs*. U.S. Government Printing Office: Washington, DC.

United States Department of Education. 1986. *What works: Rresearch about teaching and learning*. U.S. Government Printing Office: Washington, DC.

United States Department of Education. 2006. *Meeting the challenge of a changing world: Strengthening education for the 21st century*. U.S. Government Printing Office: Washington, DC.

Vitousek PM. 1994. Beyond global warming: Ecology and global change. *Ecology* **75**(7): 1861-1876.

Waters Foundation. 2003. *Systems Thinking in Schools*. http://www.watersfoundation.org. August 2006.

Weilbacher M. 1997. Confronting the enemy: Why our students are environmentally illiterate. *Clearing* **97** : 17-19.

Weinstein G, Fantini MD. 1971. *Toward Humanistic Education: A Curriculum of Affect*. Praeger: New York.

White P. 1992. The anthropomorphic machine: Causal order in nature and the world view of common sense. *British Journal of Psychology* **82**: 61-96.

Wiener N. 1954. *The Human Use of Human Beings: Cybernetics and Society*. Double Day & Company, Inc.: Garden City, NY.

Willis G, Shubert W, Bullough RV, Kridel C, Holton JT, Eds. 1993. *The American Curriculum: A Documented History*. Greenwood Press: Westport, CT.

Wright J, Muller-Landau H. 2006. The future of tropical forest species. *Biotropica* **38**(3): 287-301.

The Yale report. 1828. *The American Curriculum: A Documentary History*. Willis, George, Schubert, William H., Bullough, Robert V., Kridel, Craig, and Holton, John T., Ed. Greenwood Press: Westport, CT.

BIOGRAPHICAL SKETCH

Richard Plate was born and raised in Winter Park, Florida. He has proven consistently unable to stay out of school. After receiving a BS in chemical engineering from Clemson University in 1994 he was out for only two years before returning to Clemson and earning an MA in English in 1998. He then managed to stay away for three years before the relapse that resulted in this dissertation. When not being a student, he whiled away his time conducting water quality research in the South Pacific, freelance writing for publications nobody's ever heard of, and introducing tenth-graders to the wonders of Shakespeare and ancient mythology.