A Systems Approach

to

Science Education

Report presented by
Dr. Elizabeth S. Charles (principal researcher) and
Dr. Sylvia T. d'Apollonia (co-researcher), Dawson College.

This research project was made possible thanks to funding from the Programme d'aide à la recherche sur l'enseignement et l'apprentissage (PAREA).

PAREA project title: *A Systems Approach to Science Education*

PAREA funding from: *August 2001 – June 2003*

Report submitted: *July 2003*

Institution: *Dawson College*
3040 Sherbrooke West
Montreal, QC., H3Z 1A4

Researchers: *Dr. Elizabeth S. Charles (principal researcher)*
Dr. Sylvia T. d'Apollonia (co-researcher & project mentor)

To contact researchers: *echarles@place.dawsoncollege.qc.ca*
sdapollonia@place.dawsoncollege.qc.ca

Research project coordinator: *Mr. Bruno Geslain, Coordinator Research & Development, Dawson College*

Bibliothèque nationale du Québec

Bibliothèque nationale du Canada

ISBN number: 1-55016-169-5

ACKNOWLEDGEMENTS

We would like to thank the following individuals for their support and contribution of time and expertise towards the completion of this research project.

PAREA
François Hardy, Director of PAREA, PART, and PSCC programs - DGER

Coordination and management of research program at Dawson College:
Bruno Geslain
Suzanne Prevost

Administrative level support at Dawson College:
Neville Gurudata, Academic Dean
Michelle Serano, Director of Student Services
Ron Spivock, Dean of Technical Programs
Graeme Welch, Dean of Pre-University Programs

Technical support:
Jean-Claude Duval
Edie Howell
Jim McGill

Faculty support:
Jaleel Ali
Grant Brown
Hélène Nadeau
Claire Sergant
Peggy Simpson

Research assistant:
Charles Stevens

Thesis Committee,
Concordia University:
Gary M. Boyd, supervisor
Robert Bernard
Richard Schmid
Stephen Shaw
Michael Jacobson, external
Calvin Kalman, external

Others:
Martin Coles
Allison Hadwin
Anne Brown-MacDougall
Robert Mansoor

Finally, we would like to thank the incredible group of students that participated in this research project. These hard working and dedicated young people are an inspiration to us all and are testament to the depth of learning that is possible at the college level.

ABSTRACT

This investigation focuses on the changes in students' conceptual structures of complex systems (e.g., emergence, dynamic equilibrium, nonlinearity) after using the computer-modeling tool StarLogo with the support of cognitive coaching (ontological training). Three simulations were used: Slime, FreeGas, and Wolf Sheep Predation. Two questions were addressed: *Do students' explanatory frameworks of scientific phenomena (ontological beliefs) change as a consequence of ontological training?* And *What Complex Systems concepts do students acquire during the instructional activities?*

The results of this study suggest that the ontological training facilitated the creation of emergent framework mental models (EFMMs). This conclusion is supported by evidence that five of the nine students learned six Complex-Systems concepts: multiple-levels of organization, local interactions, probabilistic causes, open systems, random behavior, and tags. However, not all simulations were equally able to facilitate understanding of the Complex Systems concepts and none facilitated an understanding of nonlinearity, diversity and simple rules.

These results support educational efforts to use modeling tools to reify and build more global representational frameworks that are consistent with science experts.

TABLE OF CONTENTS

LIST OF FIGURES vi
LIST OF TABLES vii

INTRODUCTION 3
1.1 The Problem 5
1.2 Theoretical Foundation 7
1.2.1 Complex Systems as a Way of Thinking 8
1.3 Purpose and Significance of the Study 10

LITERATURE REVIEW 11
2.1 Conceptual Change Theories 12
2.1.1 Dissatisfaction and Knowledge Replacement Theories 14
2.1.2 Knowledge Restructuring Theories 15
2.2 Chi's Ontological Reassignment Theory 20
2.2.1 Evolution of Ontologically-based Conceptual Change Theory 20
2.2.2 Recent Refinements to the Theory 21
2.2.3 Summary of the Ontological Reassignment Theory 22
2.3 Ontologically-based Misconceptions 23
2.3.1 Misconceptions on the Topic of Evolution 23
2.3.2 Chemical Equilibrium, Diffusion, and Osmosis Misconceptions 24
2.3.3 Deterministic Causality Misconceptions 24
2.3.4 A Global Perspective on Ontologically-based Misconception 26
2.3.5 Summary of Ontologically-based Misconceptions 26
2.4 The Alternative Causal Ontological Category 27
2.4.1 Emergent Processes 26
2.5 Challenges of Teaching and Learning Complex Emergent Systems 30
2.5.1 Overcoming Lack of Awareness of Emergent Processes 31
2.5.2 Summary of Identified Challenges to Learning About Emergent Processes 32
2.6 The Status of Research on the Teaching of Emergence and Complexity 33
2.6.1 An Empirical Study on Teaching Emergence 33
2.6.2 Using Models to Teach about Emergent Causation 34
2.6.3 Using Physical Models to Teach About Complex Systems 35
2.6.4 Computer Based Modeling Environments 35
2.6.5 Emergent Process Models – *Life* 36
2.7 Main Research Questions 37

METHODS
3.1 Research Design 38
3.1.1.Participants 39
3.1.2 Instructional Intervention 39
3.1.3 Case Study Design 40

3.2 Data Collection, Coding, and Analysis of Three Outcome Measures 42
3.2.1 Explanatory Framework (Ontological Beliefs) Taxonomy 46
3.2.2 Conceptual Understanding of Complex Systems Taxonomy 49
3.2.3 Conceptual Development of Students' Systems Thinking 53

RESULTS
4.1 Question 1: Changes in Explanatory Frameworks 57
4.2 Question 2: Conceptual Understanding of Complex Systems 60
4.2.1 Individual Differences Among Students 63
4.2.2 Acquisition of Concept of Multiple Levels of Organization 66
4.2.3 Acquisition of Concept of Local Interactions 68
4.3 Question 3: Development of Students' Systems Thinking 71
4.3.1 Group 1: "Sophisticated" Understanding of Complex Systems 72
4.3.2 Group 2: "Moderate" Understanding of Complex Systems 77
4.3.3 Group 3: "Novice" Understanding of Complex Systems 82
4.3.4 Summary on Results of Concept Mapping Activity 87

DISCUSSION
5.1 Limitations to Understanding the Ontological Category of Emergence 88
5.1.1 Multiple Levels of Organization 89
5.1.2 Local Interactions 91
5.1.3 Nonlinearity and Random Behavior 92
5.2 Triangulation of Data Sources 95
5.3 Educational Implications 97
5.4 Future Directions 98

CONCLUSION 99

REFERENCES 100

APPENDICES
A. Data Collection Instruments. 111
B. Coding Templates – EMMF answers & CWWM answers. 113
C. Procedure Used in Constructing Coding Taxonomies. 117
D. Examples of Metacognitive Discourse 123

LIST OF FIGURES

Figure 1. Expert's concept map of complex systems concepts. 56
Figure 2. Advance-learner concept maps of complex systems concepts. 56
Figure 3. Changes in explanatory frameworks or ontological beliefs. 58
Figure 4. Student's understanding of Complex Systems concepts over three simulations. 64
Figure 5. Changes in the number of statements on multiple levels of organization. 68
Figure 6. Changes in the number of statements made by individual students on local interactions. 70

Figure 7. Greg's Concept Map - Session 2 & 6. 73
Figure 8. Mitch's concept map - Session 2 & 6. 74
Figure 9. Sam's concept map - Session 2 & 6. 76
Figure 10. Walter's concept map - Session 2 & 6. 79
Figure 11. Sidney's concept map - Session 2 & 6. 80
Figure 12. Norman's Concept Map - Session 2 & 6. 81
Figure 13. Penny's Concept Map - Session 2 & 6. 85
Figure 14. Emile's Concept Map - Session 2 & 6. 86

LIST OF TABLES

Table 1. Comparisons of conceptual change models. 13
Table 2. Data analysis matrix. 41
Table 3. Jacobson's "component beliefs" categories (adapted Jacobson, 2000). 46
Table 4. Second phase developments in OMMT coding schema. 47
Table 5. Final OMMT schema used to the analyze data produced in Study 1 & 2. 51
Table 6. Gains in students' mental models (n = 9). 58
Table 7. Meaning of changes to mental model categories. 59
Table 8. Complex systems concepts identified by student and reported by session. 61
Table 9. Distribution of Responses within Complex Systems Taxonomy (CST) for each simulation. 62
Table 10. Relative number of statements made by each student coded into Complex Systems concepts over the three simulations. 65
Table 11. Scoring of Concept maps on criteria 1 & 2 over time. 71
Table 12. Pearson product correlations between different data sources. 96
Table 13. Classification of case study students across the three data sources. 97
Table 14. Finding themes within the data and testing of possible links. 120

Dedicated to

All who strive to better understand the process of learning
and those who support this effort.

INTRODUCTION

One of the current trends in science education is to look at the generalities that are present across disciplines. When scientific theories are investigated at the philosophical level it is possible to abstract a general conceptual framework from the specifics that differentiate them (Auyang, 1998). These general concepts (i.e., thing, event, process, causality, relationship, part and whole, space and time, etc.) may also be referred to as explanatory frameworks used to interpret and explain the world. Furthermore, they appear to be the component characteristics of distinct ontological categories (Chi, Slotta, & deLeeuw, 1994) that allow us to share ideas and build theories. Auyang states:

> The categorical framework is general and not biased toward any theory. It abstracts from scientific theories, which in turn abstract from the specifics of wide ranges of phenomena. The shared categorical framework of different theories reveals both the general way of our thinking and a general commonality among the topics of the theories thus our consideration must include the phenomena the sciences try to understand. (p. 9)

But what happens when we do not share similar explanations for a major class of phenomena? Or worst yet, we do not even “see” similar categorical frameworks extending across various phenomena? It is theorized that this is precisely why many science concepts are difficult for the novice to grasp. In science, many underlying structural and process attributes are not consistent with the surface features of the phenomena. Novice learners tend to build explanations (mental models) based on surface features and their intuitive, naïve interpretations therefore lead to incorrect conclusions and misconceptions (Glynn & Duit, 1995; Chi, Feltovich, & Glaser, 1981; Chi, Slotta, & deLeeuw, 1994; Vosniadou, 1994).

If this is the case, what if we were to take a different approach: identify common underlying structural and process attributes and teach these instead? Would that facilitate deeper level understanding? These are the fundamental questions of this present study. We argue that teaching students (novice learners) certain principles of complex systems thinking is possible and that students are then able to change certain aspects of their explanatory frameworks used in solving both near to the model and far from the model problems. We also argue that the behaviors of complex systems, as demonstrated in the multi-agent modeling language StarLogo, are consistent with identified dimensions of the category of "emergence" type ontological beliefs (referred to as "component beliefs" in Jacobson, 2000) and therefore offer affordances for learning about underlying emergent causal processes.

Operational Definitions

Before moving forward it is necessary to clarify the intended meaning of the key concepts used in this study. To begin, the term "emergent processes" or "emergence" is used primarily to describe the behavior that arises from a collective (meta-agent) composed of a large number of smaller parts (the agents) that do not themselves exhibit behavior at all like that of the whole. Emergence is also described as the ability to generate complex behavior (complexity) from a small set of laws or rules. To illustrate this description, patterns generated by birds flocking, ant trails, the "human wave', are all examples of emergent processes. Another less visual example of emergence is a board game. Although there are many built-in constraints imposed by the rules, the outcomes are too numerous to describe (Holland, 1998). Thus emergence, is defined as: *a phenomenon which relies on the interactions of multiple agents, all operating under the same constraints (rules) without centralized control, yet affected by probabilistic causes and feedback loops that generate nonlinear effects creating dynamic self-organizing systems behaviors*.

Many, but not all, complex systems exhibit emergent behaviors. It should be noted that in this study we were primarily interested in complex systems thinking that relates to the understanding of emergent causal processes. Therefore, only references to emergent causal behaviors were identified and analyzed from the students' interactions with the StarLogo models.

The definition of the term ontological category as used in this paper is: *a personally held "theory" or "belief" associated with what the world is and what it contains* (Flood & Carson, 1993). It can be said also that it is "our way of explaining the world"; hence it contains basic axioms, and a set of causal predicates that can be judged as true or false (Keil, 1979).

The phrase "emergent causal framework ontology" refers to the ontological category that offers an explanation for the behaviors of the particular types of phenomena that exhibit the behaviors described above (i.e., aggregating, decentralized control, nonlinear effects, random actions, probabilistic causes, and dynamic self-organizing nature). Although there may be other less liberal definitions of the emergent causal ontological category (see Perkins & Grotzer, 2000) this more inclusive definition allows us to bring greater consistency between conceptual change research (Chi 2000b) and the expanding literature of complexity research (e.g., Holland, 1995; 1998; Jacobson, 2000; Wilensky 1999). Later in this paper the development of the ontological category related to emergent causal explanations will be expanded upon, as well as the mental representations that are evidence of these internally held explanatory frameworks.

A final point of clarification: the term “emergent processes” or “emergence” should not be confused with either everyday usage or with the meaning as seen in qualitative data analyses where categories and themes “emerge” from the data. Thus the reader should pay attention to the context in order to ensure they are correctly interpreting the intended meaning for this terminology.

1.1 The Problem

The literature of schema-based learning theories describes three types of learning: accretion, tuning and cognitive restructuring (Rumelhart & Norman, 1976). Research conducted in the area of cognitive restructuring generally is assembled under the heading of conceptual change. Growing interest in this field led to the identification of difficulties in learning key science topics such as electricity in physics (Chi, Feltovich, & Glaser, 1981; White, 1993), gas laws and equilibrium in chemistry (Wilson, 1998), and in the biological sciences such concepts as diffusion, osmosis (Odom, 1995; Settlage, 1994), and evolution (Anderson & Bishop 1986; Brumby, 1984; Jacobson & Archodidou, 2000).

It is very difficult to achieve cognitive restructuring, generally referred to as conceptual change. Thus far, research has shown that the presentation of anomalous data – information that contradicts the pre-instructional beliefs and theories — is generally met with resistance from the learner and seldom leads to conceptual change (Chinn & Brewer, 1993; Chinn & Brewer, 1998). There are several researchers who have followed this route and achieved differential levels of success (Tao & Gunstone, 1999; Chan, Burtis & Bereiter, 1997; Windschitl & Andre, 1998). They have combined constructivist instructional strategies, such as collaborative learning, computer-based instruction, knowledge building activities, metacognitive prompts, multiple analogies, in order to facilitate the acceptance of the anomalous data. However, these studies have not attempted to build a theory of conceptual change. Further, they have not addressed the underlying cause of the misconceptions. Chinn and Brewer (1993) propose that the crux of the problem is the learner's efforts to coordinate theory and data. These authors offer four characteristics that may account for the different responses to anomalous data. They are as follows:

- Entrenchment of prior theory
- Ontological beliefs
- Epistemological commitments
- Background knowledge

Chinn and Brewer (1993) also tell us that in the case of robust misconceptions, these four characteristics may not play an equal role. Intuitive "naïve" beliefs about the nature of existence and the fundamental categories and properties of the world (ontological beliefs), and beliefs about knowledge and how it is acquired (epistemological beliefs), may be deeply intertwined with who we are, as well as how and what we can learn. If we accept their argument, then these beliefs are likely associated with hopes and fears and therefore are rigorously defended and very resistant to change.

Conceptual change models fall into two primary groups, the more conventional view known as an accommodation model, posited by Piaget, and elaborated on by Strike and Posner (1985, 1992) who consider the conceptual ecology of the learner but assert that, through reason, the more fruitful explanation will be adopted. The other camp takes a more structural approach, positing that it is the very nature of the explanation, the underlying beliefs of causation that need to be addressed. Within these models are: (1) Vosniadou's "framework theories" (e.g., Vosniadou & Brewer, 1994), (2) diSessa's "causal net" (diSessa & Sherin, 1998), and (3) Chi's "ontological beliefs" (Chi et al. 1994). Although these researchers disagree on several fundamental points related to how coherent or fragmented these naïve "theories" or beliefs are, they agree that these beliefs need to be altered in order to repair and/or remove misconceptions.

Arguably the problem of robust misconceptions is an important problem to be solved, furthermore, a general theory of conceptual change is required. In order to move the debate forward, it is necessary to empirically as well as theoretically explore the assertions of these models. This present study addresses this need.

1.2 Theoretical Foundation

This study took as its starting point the conceptual change theory proposed by Chi and her colleagues (Chi, 1993; Chi & Slotta, 1996; Chi et al., 1994; Ferrari & Chi, 1998; Slotta & Chi, 1996, 1999; Slotta, Chi & Joram, 1995). The basic assumption of their theory is that all conceptions are classified into ontological categories — ordered hierarchical trees of super-ordinate and subordinate systems — based on attributes that are perceived or suggested to the learner. These schema-like associations act as facilitators or inhibitors of future transfer of knowledge and are part of general accretion and tuning.

Chi's theory is intended to explain concepts that, in the case of science education, fall within the ontological category of "processes". Within this category there are "event-type" processes and "emergent-type" processes. It is hypothesized that most of the misconceptions occur when concepts (e.g., electricity, osmosis, diffusion, equilibrium, evolution), which scientifically speaking, belong to the emergent process ontology, are assigned to other ontological categories (e.g., heat, electricity to the category of "material substances"; or evolution and diffusion to the category of "events"). Thus it is hypothesized that novices, unlike experts, assign concepts to ontological categories that are unable to support explanations of the phenomena, thereby acquiring robust misconceptions and flawed knowledge acquisition. Slotta and Chi (1999) state, "once an ontological commitment is made with respect to a concept, it is difficult for this to be undone" (p.8). The basic assumptions of Chi's conceptual change theory may thus be summarized as follows:

a) Concepts acquire membership in ontological categories through common language (predicates).
b) Concepts are assigned to an ontological tree in a hierarchical structure — therefore the structure of knowledge in categories is hierarchical.
c) The teaching of a new ontological category is possible.
d) The reassignment of an entity to another ontological category is necessary and possible for conceptual change and understanding.
e) Ontological attributes are distinct for members of each ontological tree. There are differences in the attributes of entities that belong to different trees.

f) Novices, more often than not, place entities into ontological categories based on surface level features.

g) Several concepts display contradictions between their surface features and their deep level features. On the surface, the attributes of these concepts resemble one type of ontological category while their veridical attributes belong to another ontological category. Chi et al. (1994) suggest that many scientific concepts require conceptual change across trees and that is why they are difficult to learn.

The actual mechanisms of assignment to ontological categories are a very important part of the discussion and understanding of Chi's conceptual change theory. It is postulated that explanatory frameworks are used to organize and express the ontological commitments present in the learner's cognitive structure (Slotta & Chi, 1999). These authors argue that explanatory frameworks are the key to operationalizing the learner's ontological commitments, and identifying the ontological boundaries between two explanatory frameworks is a means to assess conceptual change . Hence, reassignment of concepts from one ontological category to another is taken to entail learning a new explanatory framework. This may require learning new terminology, acquiring new mental models through expository and discovery learning, and may even demand new attitudes and values (e.g., epistemological beliefs, control beliefs, emotional loadings, and changed motivation).

1.2.1 Complex Systems as a Way of Thinking

Mitchel Resnick, Uri Wilensky, and Walter Stroup have championed research pertaining to the use of complex systems as a better way of thinking about science. These authors have used the computer environment and the power of multi-agent modeling language (MAML) programming to create simulations that demonstrate characteristics of complex systems that challenge naïve ontological beliefs about centralized versus decentralized control, determinacy versus randomness , order versus chaos. "In the minds of many, the study of complexity is not just a new science, but a new way of thinking about all science, a fundamental shift from the paradigms that have dominated scientific thinking for the past 300 years" (Resnick & Wilensky, 1997, p. 4). Initial studies conducted by Resnick (1994) and Wilensky (1995) tell us that the use of particular types of simulations can afford understanding of specific aspects of complexity – knowledge of the process of emergence and the subsequent development of non-isomorphic levels of organization. They have demonstrated that the use of Starlogo simulations is a powerful means of destabilizing simplistic entrenched conceptions and of facilitating multi-level thinking. Their work has focused mainly on young learners with very few other empirical studies to date.

Jacobson's (2000) most recent work has further explored the relationship between complex systems concepts and conceptual beliefs. He has demonstrated that novice learners', when solving specific types of problems, use “component beliefs” (ontological and epistemological) that are correlated with what he calls “clockwork” theories which are reductive and influenced by a Newtonian view of science. On the other hand, experts, in solving the same problems, used component beliefs that were correlated with “emergent” theories, equated to complex systems concepts. Results from Jacobson’s study were obtained from a small sample and are non-parametric provisional findings; however, they indicate a significant qualitative difference between expert and novice thinking when solving emergent framework questions. Hence, this present study aspires to contribute to the body of literature that contends there is a relationship between component beliefs and conceptual change, and complex systems thinking.

In addition, there is a body of literature that argues for the use of elements of complexity theory in the classroom. Boyd (1997) suggests that it is possible to introduce elements of “cybersystemics” into the regular curriculum. Others such as Auyang (1997), Bar-Yam (1997), Kaput, Bar-Yam, and Jacobson (1999) contend that complex systems may function as a unifying and cross-disciplinary theme. In fact, at the most recent New England Complex Systems Institute annual conference, Jacobson, Jakobsson, Lemke, and Wilensky (2002) challenged the science education community to explore the potential of using complex systems ideas in the classroom. They stated: “the conceptual basis of complex systems ideas reflects a change in perspective about our world that is important for students to develop, as it corresponds to the scientific environment that will exist when they graduate. This perspective emphasizes both the limits of predictability as well as the possibility of understanding indirect consequences of actions taken, both positive and negative, through modeling the interdependence of our world” (p.2).

1.3 Purposes and Significance of the Study

Hence, this current study will focus directly on how the properties of complex systems such as emergence, non-linear behavior, and probabilistic behaviors, which appear to be particularly challenging to "naïve" learners, are constructed when explicitly attended to in instructional activities. In so doing, we addressed the following three questions:

1. Do student’s explanatory frameworks of scientific phenomena (ontological beliefs) change as a consequence of an instructional intervention utilizing simulations of complex systems (i.e., StarLogo) and supported by cognitive scaffolding?
2. What Complex Systems concepts do students acquire during the instructional activities?
3. What is the development of students’ systems thinking?

LITERATURE REVIEW

> Concepts are construed as intrinsically relational sorts of things. They are not isolated entities connected only in the service of propositions. No individual concept can be understood without some understanding of how it relates to other concepts. Concepts are not mere probabilistic distributions of features or properties, or passive reflections of feature frequencies and correlations in the world; nor are they simple lists of necessary and sufficient features. They are mostly about things in the world, however, and bear nonarbitrary relations to features frequencies and correlations, as well as providing explanations of those frequencies and correlations. If it is the nature of concepts to provide such explanations, they can be considered to embody systematic sets of beliefs – beliefs that may be largely causal in nature. (Keil, 1989, p.1).

Theories of Concepts

Our ability to evaluate the relative merit of different theoretical positions on conceptual change requires that we appreciate several basic aspects of concepts and concept formation. Conceptual structure as proposed by the "theory-based" view[1] of concept formation pose two major concerns that impact on conceptual change research. First, because of the principle assumption that concepts are embedded in a network, the implications are that change to one component may involve a change to other components and possibly the entire system. Hence, it is reasonable to posit that the resulting cognitive behavior is likely to exhibit nonlinear (defined as non-additive) and possibly emergent characteristics; a view that appears to be supported by Limón (2001). Second, and more importantly perhaps, is that causal explanations (theories) will be a focal determinant of change in concept formation and categorization process.

Conceptual change theories that focus on multiple causal mechanisms are the most promising as explanations of the process of change. Supporting this conjecture, Ahn (1998) claims that there is reason to believe that self-explanations of causes are considerably more important than explaining effects in the process of categorization. While Keil (1989) makes an even stronger case for the role of causation, suggesting that explanations of causal relationships generate beliefs. Addressing the differences between novice and experts, he states that: "it

---

[1] The literature describes four major schools of thought on concepts and concept formation:

1. Classical Logical Definitional theories
2. Prototype Probabilistic theories
3. Associative Theory-based theory
4. Atomism theory

seems almost certain that a host of interconnected beliefs about the mechanisms of objects underlies and constrains the novice's choice of problem groups… [While experts] have probably shifted away from attempting to characterize the problem space in terms of definitions-like rules and instead have incorporated a far more complex set of intricate [co-] causal relationships more along the lines of the homeostasis model [e.g., the feedback control]" (p. 261).

2.1 Conceptual Change Theories

In reviewing the conceptual change literature we selected three of the premier models that hold the most promise for a genuine theory of conceptual change. Although these models differ on several major assumptions, there are equally several specific dimensions where they support each other in a principled manner (see Table 1).

General Background

The foundations of most conceptual change theories include a Piagetian motivational theory of cognitive development. Piaget (1975) proposed that disequilibrium, dissatisfaction, or discord must be created within the child between their initial conception and the to-be-learnt one. The attempt to resolve this cognitive conflict results in the processes of "assimilation" or "accommodation"[2] of the new idea. This notion of dissatisfaction is at the base of several early models of conceptual change. However, some studies in the field have moved away from theories of cognitive conflict to theories of knowledge restructuring (e.g., Chi et al., 1994; Vosniadou, 1994).

---

[2] Assimilation refers to situations in which the learners' existing theories allow them to explain new situations; whereas accommodation describes situations in which the learners' theories cannot account for or explain the new phenomena therefore must be revised, reorganized, or replaced.

Table 1. Comparisons of Conceptual Change Models

Model	Intuitive theories	Misconceptions created by	Conceptual change occurs when	Techniques for conceptual change	Process
Strike & Posner 1985; 1992; Thagard 1987?	Many factors conceptual ecology some have explanatory coherence.	(Does not state specifically)	Confrontation or competition between two incompatible but equally well organized theories. Knowledge replacement models (other example is Thagard, 1992) Other factors: conceptual ecology including motivation and goals.	1. Instruction of alternative information. 2. Cognitive artifacts used to achieve change: - analogies - metaphors 3. Extensive use of constructivist techniques.	Four stage process: 1. Minimal understanding of new conception 2. Dissatisfaction with existing conceptions 3. Plausibility of new conception 4. Fruitfulness of new conception
Vosniadou 1994; 1998	Coherent theories: a) Framework theory b) Specific subject dependent theory	The synthetic stage where new information is added to existing initial model. 1. Initial stage 2. Synthetic stage 3. Scientific stage	Realization that new information is inconsistent with specific theories or framework theories. Restructuring is therefore a continuous gradual enrichment.	Instruction of scientific information.	Two phased process: 1. requires metaconceptual awareness. 2. change in specific beliefs and framework Two types of Concept Change: 1) spontaneous, without need of instruction. 2) instructionally based. This process is a slow revision and gradual incorporation of science explanation.
diSessa 1993; 1998	Fragmented knowledge structures in the form of p-prims.	(Does not state specifically)	Slow refinement & restructuring of intuitive knowledge structures, p-prims, to formal principles of science. Shift in way of "seeing" which is a separate change in *readout* strategies and *causal net*.	Instruction of alternative causal net and readout strategies.	Iterative process of changes in: 1. Readout strategies 2. Causal net
Chi 1993; 1994; 1997; 1999; 2000	Coherent knowledge structures but not at the level of theory, rather organized as a schema or frame.	Assignment of concept to the incorrect ontological tree.	Reassignment of concept to different ontological tree. Change in preexisting conception.	Instruction of alternative ontological process category – specifically "emergent" processes.	Shift in reassignment to correct ontological category.

2.1.1 Dissatisfaction and Knowledge Replacement Theories

Strike and Posner's (1985) model of conceptual change expanded on Piaget's ideas of conceptual change through accommodation and assimilation. These authors contended that accommodation requires taking into account the alternative conception and comparing it to the existing conception; if the plausibility status of the alternative conception reaches a level that exceeds that of the original conception, then there will be a change. Four stages were identified: (1) dissatisfaction with original conception, (2) intelligible replacement conception, (3) plausibility of new conception, and finally (4) fruitfulness of new conception leading to replacement. A central claim of the conceptual change theory proposed by Strike and Posner (1985) is that new conceptions are understood, judged, acquired, or rejected in a conceptual context – a conceptual ecology. Hence, old and new concepts coexist. Cautioning that there were many factors other than the conceptions themselves that may affect status, they suggested that the learner would experience stops, starts, and even retreats along the road to change.

In their revised discussion of conceptual change, Strike and Posner (1992) extended the role of the conceptual ecology. Proposing that misconceptions are not the product of clearly articulated beliefs, rather, they are artifacts of deeply entrenched problems in the conceptual 'ecology'. Raising the issue of stability, they suggested that misconceptions may be weakly formed, temporal and not consistent; in fact, they may be influenced by the conceptual ecology. Addressing the issue of conceptual structure they draw attention to the systemic nature of the conceptual network.

Importance of this model. The importance of Strike and Posner's model is twofold: (1) its attention to factors such as motives and goals that influence the learners' conceptual ecology, and (2) its applicability to classroom instruction. It has been the cornerstone of most conceptual change instructional interventions, however these studies have produced some equivocal results (e.g., Chan, Burtis & Bereiter, 1997; Champagne, Gunstone & Klopfer, 1985; Jensen & Finley, 1996; Limón & Carretero, 1997). Because this model does not adequately address how to build an alternative conception, it falls outside the focus of this present study and is set aside for the moment. Perhaps, when we have a better understanding of the processes and mechanisms of conceptual change, this contribution to a global conceptual change theory should be revisited for its perspective on issues of ecological accommodation.

2.1.2 Knowledge Restructuring Theories

If concepts are indeed embedded in stable complex networks of other concepts that represent naïve or personal theories, then conceptual change will be a formable task. If in addition, these theories are held together by causal self-explanations composed of the most basic units of our thinking – ontological and/or epistemological beliefs – then how do we start to unravel the problem of deeply held misconceptions that are ubiquitous in science learning (see, for example, Driver, 1995 or Pfundt & Duit, 1994)? Consider the following theories as steps toward clarifying a possible integrated causal approach to conceptual change.

Vosniadou – Framework and Specific Theories

Vosniadou (1994) argues that concepts are entrenched and constrained within a larger theoretical structure. This author identifies two levels of theories that control the learners' beliefs, naïve *framework theories* and various *specific theories*. Vosniadou proposes that the learner's framework theory is not available to his conscious awareness; nonetheless, this theory constrains the process of acquiring veridical knowledge about the physical world. These theories are a function of ontological and epistemological presuppositions. Specific theories on the other hand, are consciously accessible, exist within a domain and consist of a set of interrelated propositions that describe the observed behavior of physical objects. That is, the specific theory is based on the individual's observations, as well as the instructional information, and it is developed within the constraints of the presuppositions of their framework theory. These two classes of theories come together to create the mental model, the lens, through which the learner builds causal explanations of the world.

Definition of conceptual change. Vosniadou (1994) identifies two kinds of conceptual change: *enrichment* and *revision*. The former is described as the simple addition of new information to existing knowledge, and achieved through the process of accretion. The latter is considered conceptual change and viewed as a substantial change that is realized by the learner when new information is inconsistent with specific theories or framework theories. She posits that inconsistencies between new information and framework theories are more difficult to resolve than inconsistencies with specific theories.

Difficulties in removing misconceptions. Vosniadou suggests that conceptual change is difficult because framework theories are coherent systems of explanations that are based on everyday experiences and grounded in years of confirmation. Additionally, because these are ontologically and epistemologically based, a shift in any of these beliefs will create a shift in the entire system of the framework theory and all the other knowledge built upon it. This assertion is similar to implications of Strike and Posner's conceptual ecology.

Failure to learn certain concepts has been attributed to inconsistencies between the to-be-learnt knowledge and framework theories. These occur when children attempt to add information to the false existing mental structure. The author describes inert knowledge as the product of inconsistent information being stored in separate microstructures and used only in particular situations. Whereas, misconceptions are the result of learners trying to reconcile the inconsistent pieces of information and in the process produce *synthetic mental models*[3]. We contend that this attempt to account for anomalous data draws this part of Vosniadou's model closer to the cognitive conflict approach. Therefore, her model may be viewed as a bridge between Strike and Posner's model and those that will be described next.

Importance of the model. Vosniadou and Brewer's (1994) empirical findings suggest: (1) there is a sequence in which concepts are acquired in a conceptual domain; and, (2) that the importance of mental models is a constraint on the knowledge acquisition process. These findings have given rise to their theoretical supposition that conceptual change is gradual and will give rise to misconceptions. They also suggest that there are developmentally distinct stages in conceptual change: (1) initial mental model, (2) synthetic mental model – learner attempts to reconcile the science model with initial model, and (3) scientific mental model. Another group of researchers, Jacobson and Archodidou (2000), have successfully identified these developmental stages in their study of conceptual change instruction on the topic of evolution.

Recently, Vosniadou and Ioannides (1998) have made two major refinements to the original model. Firstly, they have identified distinctions between types of conceptual change suggesting that conceptual change can be: (1) spontaneous, or (2) instructionally based. The former type is a change resulting from enriched observations in social learning context without

[3] According to Vosniadou and Brewer (1994), *synthetic mental models* are likely to be formed when the knowledge acquisition process requires revision of framework theories based thus are part of the presuppositions constructed on our interpretations of everyday experiences. Synthetic models function as intermediary steps in the conceptual change process from an initial intuitive model to the scientifically culturally accepted one.

formal science instruction. Examples of this would be language learning as a result of socialization with adults and older children as a child matures. The latter is a result of formal instruction that requires the building of synthetic models in an effort to reconcile science instruction into existing theories (e.g., understanding of astronomical processes).

Secondly, they have elaborated on Vosniadou's (1994) original assertions regarding the refinement process. The role played by *metaconceptual awareness* has been strengthened and refinement is viewed as the development of "theoretical frameworks with greater systematicity, coherence, and explanatory power [i.e., more scientific]" (Vosniadou & Ioannides, 1998, p. 1222). This feature is an important contribution to the development of the prescriptive side of the conceptual change debate. Additionally as will be shown, it is also a consistent theme between models, although it may be argued that Vosniadou makes the biggest commitment to its importance.

<u>DiSessa's and Sherin's Model</u>

Along the continuum of conceptual change models, the work of diSessa and Sherin (1998) is positioned closer to that of Chi and her colleagues. Similar to Chi, they too focus on the deeper issues of process and mechanism of concept formation and concept change. In fact, it may even be argued that theirs is a fuller model of concept formation. we will briefly describe this model with a focus on its assertions regarding the processes of what they call the *coordination classes*, a major structural component in concept change.

<u>The basic assumptions of the model.</u> Based on the research supporting the supposition that naïve learners possess impoverished causal models for understanding physics concepts (Gentner & Stevens, 1983), diSessa (1993b) developed his model of concept formation. He identified this attribution as the "naïve sense of mechanism", suggesting that this belief of causality is composed of *phenomenological primitives* (here forward referred to as *p-prims*) which are abstracted from common experiences. P-prims are the smallest unit of particular knowledge elements[4] and indeed may generate their own self-explanations. In these cases, diSessa (1993b) states that p-prims are the intuitive equivalent of physical laws and form the bases upon which one sees and explains the world. Hence, p-prims account for structures that

[4] Not unlike "Conversation theory" concepts as described by Gordon Pasks (Boyd, 1997).

diSessa calls *causal nets*. However, p-prims are not concepts themselves, and multiple p-prims are involved in the creation of causal nets.

Causal nets may be described as approximately corresponding to what people intuitively expect of causality, which is logical given their status as composites of p-prims. In addition, in some instances, they can be interpreted to mean the reasoning strategies used to explain how some observations are related to the information at hand. In their words: “Causal nets are, roughly, our replacement for the ‘theories that lie behind observations’. Or the theories implicated in theory-based notions of categories” (diSessa & Sherin, 1998, p. 1174), hence, causal nets, may be described as the inference-based explanations used to make sense of the world, which in turn form the basis of our theories. These authors link this explanatory mechanism to concept acquisition through a structural component called a *coordination class*. In order to understand this sophisticated interwoven play of components requires some background information.

Background on coordination classes. DiSessa and Sherin (1998) first suggest that concepts are not all the same. In fact, concepts such as ‘robin’ are different from those such as ‘velocity’ or ‘force’ and require different cognitive processing. While the former requires sorting into the category-like concept of bird, the latter two fall into a special class of concepts that they refer to as *coordination classes*. These coordination classes are made of structural components that perform two distinct activities: (1) centered on gathering information through selecting what to ‘see’ (referred to as “readout strategies”), and the other, (2) based on the already mentioned causal net activity.

Part one, the *readout strategy*, or information gathering, is equated to a metaphorical ‘seeing’, and the shift in the means of seeing is considered to be the core problem of conceptual change. They state: “In many instances this seeing is a substantial accomplishment of learning and will depend only very partially on basic perceptual capabilities. In addition, these forms of seeing sometimes involve explicit strategies and extended reasoning” (diSessa& Sherin, 1998, p. 1171). We will return to this point shortly.

Elaborating on the *readout strategies* they identify two subcomponents of this phase: (1) *integration*, which refers to the fact that multiple observations or aspects may need to be coordinated so as to determine the requisite information; and, (2) *invariance* (we would suggest that it could be considered a type of concept stability), which refers to the knowledge that accomplishes the readout of information from different instances and situations, must consistently and reliably determine the same information.

Part two of the *coordination classes* process takes us back to the important explanatory mechanism, *causal nets*. Learning new science concepts therefore becomes an interlocking cognitive 'see-saw' where both *readout strategies* and the *causal net* are said to co-evolve. These authors suggest: "There should be episodes of 'conceptual bootstrapping', where causal assumptions drive the learning of new readout strategies. On other occasions, 'noticings' - for example, that something surprisingly affects something else - may drive reformulations in the causal net. In general, characteristics of one will have important influences on how the other behaves and develops" (p. 1177).

<u>Definition of conceptual change</u>. Hence diSessa and Sherin (1998) define conceptual change as involving both the separate changes in readout strategies and in the causal net. They clarify by showing an example, that it is possible that no new readout strategies are necessary in learning a new coordination class, rather existing ones come to be organized and used differently. On the causal net side, maybe the construction of a whole new causal net may be required, or an existing one may need to be developed and reorganized.

<u>Importance of the model</u>. The detail provided relating to the activity of the coordination classes is an important feature of this model. The suggestion that conceptual change is a two-part process in which conscious attending to evidence (e.g., data) followed by conscious attending to the explanations related to causation (e.g., personal theories) is a development and clarification on Vosniadou's concerns with "metaconceptual awareness". In fact, diSessa and Sherin propose that the causal net is the source of difficulty in learning school physics. Their recommendation is thus, "among other things, it [the causal net] needs to become more systematically organized. The notions of invariance and integration may play a role in the organization and selection of causal net to be used" (diSessa & Sherin, 1998, p. 1178).

It is arguable that missing from this model is the answer to the question: What kinds of changes occur in the causal net? In other words, if we are to attend to new causation what is needed to fill in this gap? To address this question we must turn to Chi's views on a theory of conceptual change for an answer.

2.2 Chi's Ontological Reassignment Theory

2.2.1 Evolution of Ontologically-based Conceptual Change Theory

The original model. Chi et al., (1994) define conceptual change as learning that changes a preexisting conception. This definition holds a basic assumption that the learner has some prior idea on knowledge of the concept, which in turn may mean that it has already been classified into a category. Therefore the meaning of a concept is determined by its category assignment and conceptual change is defined as a change in category assignment. On the other hand, the simpler process of "belief revision", according to Chi (1992, 1997), occurs when the concept just needs an adjustment to the category (an addition or deletion of information). Accordingly, the most important aspect of Chi's theory of conceptual change is this notion of re-assignment of concept from the initial category in the ontological tree to the veridical category of the tree. The way the categories in one tree differ from categories in another is embedded in their ontological attributes.

Chi's theory of conceptual change (Chi et al., 1994; Chi & Slotta, 1993) rests on three assumptions: (1) an epistemological assumption concerning the ontological assignment and beliefs about the nature of entities in the world[5], thereby defining the criterion of "different"; (2) a metaphysical assumption concerning the nature of certain scientific concepts (a position that we contend sets Chi apart from other theorists inasmuch as she takes an outside-looking-in approach that perhaps is related to her research on expertise); and, (3) a psychological

[5] There is a long tradition of theorizing about ontological categories based on predictability or use of predicates in natural languages. Keil (1979) describes the term predictability as follows: "it determines which classes of predicates can be sensibly combined with which classes of terms, and it appears to involve hierarchical organization in that a predicate P1 may be sensibly combined with a superset of the set of terms that can be sensibly combined with a predicate P2" (p. 11). Therefore an ontological category would be defined by the set of terms for which a particular set of predicates could be applied to and the statements judged to be true or false. Predicates used in such a manner that the statements cannot be judged to be veridical or fallacious, suggest that the terms do not belong to the same ontological category. An examples that are often cited would be that of colour. The predicate "is green" may apply to "the frog" (natural kind) or to "the table" (artifact) or to "the girl" (natural kind). Because these can be proven true or false, however, "is green" cannot be with "an hour" in any sensible way except metaphorically. Keil states, "a predicate spans a term if and only if that predicate-term combination makes sense and can be assigned a truth value, which can be either true or false (p. 11).

This paper will not delve further into the discussion of predictability and the assumption that it can identify different ontological categories. It will suffice to say that the study of this topic is covered by Sommer's theory of the relation between predicates and terms, which represent a class of terms leading to differential ontological categories membership (Sommer, 1963; cited in Keil, 1979, p. 15 see figure).

Thus, a primary assumption of this study is that natural kinds and artifacts belong to distinct different ontological categories, which can be identified through a term-predicate relationship. However, it would be difficult to suggest the same rigid test could hold true for the ontological category of processes, which Chi has identified. This should not negate the fact that processes fall under different rules of operation and therefore can be considered to belong to different ways of understanding the world, hence, different ontological categories. Since there is little literature in this domain, we will base this statement on the work in complexity that uses a different mode of operation for certain phenomena (referring to statements made by Jim Kaput, Uri Wilensky, and others at symposium on Complex Systems in Education, presentation at AERA, April 2002).

assumption concerning the learner's naïve conceptions and miscategorization of concepts that are revealed in a propositional context (i.e., mental models).

2.2.2 Recent Refinements to the Theory

Two major changes have appeared in Chi's description of this conceptual change theory. The first relates to the difficulty in removing misconceptions, the second to the structure of the categories.

Reconceptualization on removing of misconceptions. In her most recent publications (e.g., Chi & Roscoe, 2002), Chi clarifies her stance on the structures of concepts as embedded in naïve theories. Furthermore, she explicitly acknowledges the assumptions that naïve theories and scientific theories are often incommensurate; a statement similar to diSessa and Sherin (1998). Her most important conjecture, however, is that the major challenge in conceptual change comes from the fact that "students may lack awareness of when they need to shift [to an alternative ontological category], and may lack an alternative category to shift to" (Chi & Roscoe, 2002, p. 18). These authors postulate that in fact the lack of the scientifically appropriate category (emergent processes) prevents students from requisite recategorization: "students cannot repair misconceptions if conceptual shift is not possible. This is what makes certain misconceptions more difficult to repair than others" (p.19).

How then does one gain awareness of or access to these new categories? This is the major question posed by Chi (2000b), and Chi and Roscoe (2002); and, this is the major question that we have focused on in this study.

As described earlier, empirical studies using the anomalous data confrontation models have produced equivocal results relying on constructivist instructional strategies to bolster the potency of the treatment. Limón, (2001) states:

> Despite the positive effects we have reported, perhaps the most outstanding result of the studies using the cognitive conflict strategy is the lack of efficacy for students to achieve a strong restructuring and, consequently, a deep understanding of the new information. Sometimes, partial changes are achieved, but in some cases they disappear in a short period of time after the instructional intervention. Why are students so resistant to change even when they are aware of contradiction? Why are students able to partially modify their beliefs and theories but keep the core of their initial theory? (p. 364).

This shortcoming of confrontation is exactly what Chi and Roscoe (2002) believe is averted when conceptual change is approached from the perspective of reassignment. Perhaps an answer to Limón's first question, and maybe even the second, is indicated in their statement: "The problem is that unless students have an alternate category to reassign the concept to, such instruction [presentation of anomalous data] will not be effective" (Chi & Roscoe, 2002, p. 19).

Therefore, where do we start? Initial questions are: Do novice science students possess the suitable alternative ontological categories? If they already possess the needed alternative category, does a shift in explanation require mere facilitation or does it require knowledge reorganization? If this category does not already exist, then can we teach them about this category? How can this be accomplished, what do they learn, and how long does it take? Finally, if there is a change, is it long lasting? These are the primary questions that this study elects to address.

2.2.3 Summary of the Ontological Reassignment Theory

In 1993 Chi and Slotta compared their model of conceptual change to diSessa's (1993b) model of concept formation. They concurred then that there were several points of reconciliation between the two models. For instances, the role played by p-prims could be viewed as low-level instantiation of the category reasoning process. Continuing, they point to several specific points of agreement such as: (1) intuitive knowledge is phenomenological in the sense of it being personal empirical knowledge; (2) retrieval of intuitive knowledge is driven largely by surface features; and, (3) while intuitive knowledge is primitive and requires no explanation, it forms the basis of high-level reasoning about physical processes. However, there were and still are irreconcilable differences between diSessa and others regarding the structure of intuitive knowledge[6]. For example, Chi and Vosniadou view intuitive knowledge as coherent "theories", while diSessa's view is that intuitive knowledge is fragmented, "knowledge in pieces".

Although this difference is significant, Chi's recent focus on causation draws the two models closer. It is reasonable to therefore put forward the proposition that the "coordination class" may be a representation of an ontological category since it acts as the control mechanism regulating the two phases of concept acquisition – readout strategies (what we unconsciously choose to 'see' of the world) and causal nets (how consciously we explain what we 'see'). Furthermore, the ontological frame required to explain many scientific concepts is really an

[6] Observed in a verbal debate between Andrea diSessa and Stella Vosniadou (AERA Annual Meeting, New Orleans, 2002).

explanation or attribution of different types of "causation". The bringing together of the two theories was not the focus of this research, nonetheless, further mention will be made in the discussion section of this paper. Finally, it is plausible to propose that Vosniadou's model, although not focused on causation, suggests that some types of conceptual change may be intentional and call for effortful attending to metacognitive processes in the form of metaconceptual awareness. This possible connection between intentional learning (e.g., metacognitive and metaconceptual awareness, motivation, and epistemological beliefs) and conceptual change is a primary focus of the recent publication edited by Sinatra and Pintrich (2002). Although this aspect of conceptual change is not directly manipulated in this study, the important role of intentional reflection in the form of metacognitive activities and metaconceptual awareness was observed in the learners' during the intervention.

2.3 Ontologically-based Misconceptions

Before proceeding, it is important to evaluate Chi's assertion that many important misconceptions are ontologically based. Hence a question to be answered was: Is there empirical evidence from the literature of ontological category-like misconceptions?

There are hundreds of reported cases of different types of scientific misconceptions (Driver, 1995; Pfundt & Duit, 1994). Some are trivial in that they require simple restructuring of information, however, the ones that are discussed in conceptual change literature tend to fit with and support Chi's ontological supposition. To illustrate this point, we present a sample of studies in the following section from a variety of disciplines and sources.

2.3.1 Misconceptions on the Topic of Evolution

There is a substantial body of literature describing the difficulties involved in changing students' misconceptions in the learning of evolution (Ferrari & Chi, 1998; Jacobson & Archodidou, 2000). The problems range from the understanding of the time frames (e.g., Renner, Brumby, & Shepherd, 1981), to genetics (e.g., Demastes, Good, & Peebles, 1995; Jensen & Finley, 1996), to the distinctions between species and individuals (e.g., Hallden, 1988), the origin and survival of new traits, the role of variation within a population, and evolution as the changing proportion of individuals with distinct traits in a population (Bishop & Anderson, 1990;), to the explanation of spontaneous genetic mutation (Settlage, 1994), the evolutionary changes supposedly occurring as a result of need (Brumby, 1984), and finally failure to recognize that many aspects of evolution exhibit "equilibration-type" processes as opposed to "event-type" processes (Ferrari & Chi, 1998). Therefore a host of attributions ranging from those about teleological beliefs to those about isomorphic behaviors between levels is represented in these studies of difficulties in conceptualization.

2.3.2 Chemical Equilibrium, Diffusion, and Osmosis Misconceptions

The literature concerned with the instruction of chemistry has identified a persistent misconception about chemical equilibrium (e.g., Suits, 2000; Coll, R.K. & Treagust, D.F., 2002). These misconceptions appear to stem from misunderstanding of the differential levels of operations, as well as the different symbolic representations, that are discussed in the course of a normal chemistry lecture (Barnerjee, 1995). On the related topic of osmosis and diffusion, there is also evidence that similar misconceptions exist (e.g., Odom, A.L., 1995; Sanger, M. J., Brecheisen, D.M., & Hynek, B.M., 2001). Again the attributions of isomorphic behaviors between levels as well as assumed static behaviors once equilibrium is achieved are common themes. These empirical studies lend support to Chi's conjecture that there is an ontological base to this class of science misconception.

2.3.3 Deterministic Causality Misconceptions

From the literature on judgment and decision-making, evidence suggests that both adults and children exhibit difficulty reasoning about uncertainty with greater tendencies to attribute deterministic outcomes in problem solving (e.g., Shaughnessy, 1992; Tversky & Kahneman, 1974). It is arguable that the findings from studies relating to mathematics and statistics may not cross over to problems encountered in other domains of science; however, this is not the case with attributions of determinism. According to Metz (1998): "Without an understanding of randomness and probability, formal study in statistics can have little meaning, and informal [mis]interpretations of patterns and variability in the world around us will frequently result in spurious causal attributions" (p. 286). The latter assertion is precisely the one that supports this dimension to the ontological category of emergent phenomena.

Another interpretation of randomness and probability. In this study we use the terms randomness and probability in ways that may be unfamiliar to the lay person. For instances, randomness is an important behavior that accounts for much of the variety and requisite error observed an emergent phenomena (see Bar Yam, 1997). With regard to probabilistic behaviors, they occur as the interactions of the multiple agents, and systems, producing stochastic outcomes thereby making causal mechanisms more complex.

Even using these descriptions of randomness and probability there is evidence of misconceptions (Wilensky, 1993; 1995; 1997). The early work from the MIT labs (Wilensky & Resnick, 1995; Wilensky & Resnick, 1999) suggested that conceptions of randomness as being destructive are prevalent in non-scientific reasoning. Furthermore, there is abundant evidence that people have difficulty reasoning about parallelism[7] and probability. Resnick and Wilensky have thus argued, together and independently, that these misconceptions be considered the "deterministic/centralized mindset"(further elaboration to follow).

2.3.4 A Global Perspective on Ontologically-based Misconception

Spiro and his colleagues (Feltovich, Coulsen, Spiro, & Adami, 1992; Spiro, Coulson, Feltovich and Anderson, 1988; Spiro, Feltovich, Jacobson, & Coulson, 1992) have identified a range of misconceptions exhibited by medical students. One of the common attributions they identified in learners, even in advance stages of learning, was the tendency to adopt a reductive approach to problem solving; in addition to oversimplification of the subject matter. They identified three biases: "[1] *additivity bias*, in which parts of complex entities that have been studied in isolation are assumed to retain their characteristics when the parts are reintegrated into the whole from which they were drawn; [2] *discreteness bias*, in which continuously dimensioned attributes (like length) are bifurcated to their poles and continuous processes are instead segmented into discrete steps; and [3] *compartmentalization bias*, in which conceptual elements that are in reality highly interdependent are instead treated in isolation, missing important aspects of their interaction" (Spiro et al., 1992, p. 26).

If we examine the biases individually, they can be matched to the naïve attributions listed above. Hence, the "additivity bias" may also be described as attributions of isomorphic behaviors

[7] "Parallelism" refers to the simultaneous operations of multiple agents/objects all programmed to do the same thing. Complex systems operate in this manner therefore simulations are a good way to represent this type of functioning.

between levels (reductive or non-emergent behaviors). While the "discreteness bias" shows signs of attributions of static outcomes or beginning end processes. Finally the "compartmentalization bias" may be related to the global lack of awareness of emergent properties.

Latest developments. Jacobson's (2000) research identified specific categories of attributions that differentiate expert and novice reasoning about emergent phenomena. These categories will be described in the methods; however, for the moment it is sufficient to state that these categories support the data described above as well as Chi's ontological dimensions.

2.3.5 Summary of Ontologically-based Misconceptions

In summary, we argue that these studies provide some reasonable confirmation of the articulated interpretation of the ontological categories described by Chi (2000). Furthermore, although Spiro et al. (1992) research (i.e., Cognitive Flexibility Theory, CFT) addressed difficulties associated with advanced knowledge acquisition, CFT may also apply to challenges faced in the removal of robust science misconceptions. For these reasons, we borrowed from the recommendations made by that theory to design the instructional intervention used in this current study.

2.4 The Alternative Causal Ontological Category

If a possible solution to the removal of robust misconceptions is the reassignment to an alternative ontological category, then what are the problems associated with the learning of such categories? Before attempting to answer this question, however, we must first define what is meant by the ontological category of emergent causal processes, and describe why they constitute a 'different' causal explanation.

Chi and her colleagues have spent the past decade describing this category using a variety of terms: "events" (Chi, 1992), "acausal interactions" Chi and Slotta (1993), "constraint-based interactions" (Chi, Slotta & deLeeuw, 1994; Slotta, Chi & Joram, 1995; Slotta & Chi, 1996), "equilibration" processes (Chi, 1997; Ferrari & Chi, 1998; Slotta & Chi, 1999), "CDS" (Chi, 2000b), and currently "emergent" processes (Chi & Roscoe, 2002). This latest naming is the most parsimonious and consistent with the existing literature on the processes that Chi has described for years using those varied terms. Although Chi has not acknowledged the body of literature from the field of complex systems, and indeed even distances herself from it (Chi, 2000b), her latest writing (Chi & Roscoe, 2002) brings her closer than ever before to affirming a reasonable connection.

2.4.1 Emergent Processes

In order to describe the emergent processes ontological category we first turn to the literature that describes it best: that is, complexity and complex systems theory. Born out of disciplines such as biology, cybernetics, mathematics, statistical mechanics, and quantum physics, the theories related to the phenomena of complexity are undeniably daunting. Nonetheless, the ability of these theories to explain the behaviors of countless biological, chemical, physical and social interactions requires that we take a serious look at their potency as representational structures in our curricula.

Genesis of Interest in the Topic

Kauffman (1995) identifies complexity as the state at which a system of many coupled components is "orderly enough to ensure stability, yet full of flexibility and surprise" (p. 87). He continues to describe this state as one, which is just near phase transition and referred to as the "edge of chaos". One might assume from these beginnings that the study of complexity should be reserved for biologists and mathematicians; however, there is another side to this area of study. It is the conceptual side of complexity where the behaviors of countless phenomena, including social, economic, cognitive, and scientific, can be explained using the global structural features of complex systems.

Waldrop (1992), a science writer, confirms in his book entitled Complexity that at present the field is still poorly defined as researchers grapple with questions that cut across the traditional disciplines. This observation may also apply to the terms used to define the field. Some researchers refer to it as the study of complex systems, or complex adaptive systems, while others identify it as the study of self-organizing systems, and some as emergent systems. These differences should not be viewed as weaknesses, rather, as a sign of the newness of the area. Indeed, it is difficult to describe individual behaviors of complex systems because of their interconnected nature. In an attempt to keep the different terminology to a minimum, in this current study we will define the term "complex systems"[8] to refer to both complex adaptive systems as well as complex non-adaptive systems unless the adaptive nature is essential to describing the system. Additionally, the term *emergence* or *emergent properties/behaviors* of systems is used as the exemplification of the characteristic displayed by some types of complex systems as well as the product of *self-organization*.

What is emergence ? A simple explanation of emergence is a phenomenon wherein the interaction of a system's parts results in a higher order organization which behaves differently from what one could predict from knowledge of the parts alone. Hence, the commonly known cliché "the whole is greater than the sum of the parts" is an apt description.

Emergence, however, is anything but simple to describe or to understand. In an effort to explain the phenomenon, we turn to two authors, John Holland and Yaneer Bar-Yam, who each have written extensively on the subject. Holland (1998) describes emergence as patterns of interactions that persist despite a continual turnover in the constituents of the patterns. In an effort to make the concept more accessible, Holland uses a metaphor of a checkers game where the rules are invariant but the outcome of the interactions are varied and never dull, particularly in the hands of skilled players. Accordingly, it is difficult to make predictions about behaviors of emergent systems even when the rules and initial states are specified. A difficulty that is compounded when the system is composed of mechanisms that allow for adaptation and learning; that is, overt internal models with *lookahead* protocols (these well be discussed more fully later on). Nonetheless, over time both adaptive and nonadaptive emergent systems exhibit recurring patterns that are discerned by attending to specific details. Therefore, such patterns are an important property of emergent systems and can be used to characterize them without reference to underlying strata.

Bar-Yam (1997) describes emergence as the behavior that arises in the collective that is not exhibited in the behavior of the parts (nor would arise from a simple summation of behaviors). He is quick to point out that although the collective behavior is not readily understood from the behavior of the parts, this should not be taken to mean that the collective

[8] Definition of complex systems: A system is a hierarchically organized collection of a large number of coupled components defined by stated boundaries. The smallest unit of a system is referred to as an agent. Complex systems are a category of systems characterized by highly interacting individual agents operating under specified rules resulting in emergence of meta-agents and/or systems that exhibit differential behaviors to their component agents. Complex systems may be of the adaptive type (e.g., human beings, the immune system, viruses, etc.), which form internal models that learn and evolve over time or the nonadaptive type, which does not exhibit adaptive qualities (e.g., molecules, galaxies, etc.).

behavior is not “contained in the behavior of the parts if they are studied in the context in which they are found” (p.10). This subtlety leads to a distinction between two types of emergent behaviors, local emergence and global emergence. Local emergence implies that taking a small part out of a large system would result in little change to the properties of the small part or the properties of the larger system. Examples of this would be water droplets that contain the properties of water regardless of how small a quantity of water we look at (e.g., one molecule of H_20 has no fluidity). In contrast, global emergent properties invest greater interdependence of parts. For instance, an emergent traffic jam that propagates backwards despite the forward motion of the individual cars; or the parts of the brain, or a corporation that are different *in situ* compared to their isolated parts. Hence, a small part cannot be studied outside of the larger system and still exhibit the properties it has when embedded in the whole system.

Operationally defining emergence. Although Bar-Yam (1997) points to important distinctions in the phenomenon of emergence as part of the study of complexity, for the purpose of this discussion, we selected the more general description that views a system’s emergent properties as patterns or recurring structures resulting from non-linear interactions, of lower level parts (agents), governed by specific rules and relationships. These rules and relationships are the mechanisms that afford emergence, which is the resultant state of coupling all the lesser processes of self-organization/aggregation, nonlinearity, stochastic behavior, tagging/selection, flows of information. Consequently, emergence is the topmost attribute of the selected system that uses the mechanisms of self-organization[9] to produce the emergent outcome.

[9] Self-organization (as well as the process of selection) is so pervasive in nature that Kauffman (1995), one of the first scholars to write about the subject, likened it to universal laws. In his book At Home in the Universe, he builds an argument to explain how these processes made “the emergence of life well-nigh inevitable” (p.43). Waldrop (1993) describes self-organization as a process wherein “groups of agents seeking mutual accommodation and self-consistency somehow manage to transcend themselves, acquiring collective properties such as life, thought, and purpose that they might never have possessed individually” (p.11). Does this mean that self-organization is part of emergence? Yes, as agents go through the process of self-organization, there is an emergent meta-agent created as described in the previous paragraphs.

In explaining principles of self-organization particularly in developmental biology, Bar Yam touches on an essential quality in understanding the beauty of this process, that is, its economy. Self-organization is a process in which the representation describes the developmental process of formation rather than the final system itself. It is thus the creation of an algorithm from which the system is to arise. Another parsimonious property is that randomness or noise acts as a bonus to the unfolding of the algorithm in that it adds the element of chance variation without breaking the reproduction value. For further information we direct the reader to Bar-Yam’s book.

2.5 Challenges of Teaching and Learning Emergent Causal Processes

Why do concepts related to the emergent processes category prove challenging to learners? Chi (2000) suggests three possible challenges to the removal of misconceptions: (1) the nature of human cognition, (2) nature of instruction, and (3) students' lack of awareness about the nature of emergent processes.

Nature of Human Cognition

To support this conjecture, Chi cites two separate sources. First, she turns to Wilson and Keil (2000) who posit that humans are predisposed to think simplistically in causal terms (see p.5 this document, as well as diSessa's explanation of the role of causal nets, for other support of this argument). Second, she suggests that in normal cognitive development and learning, rarely are ontological shifts required. In fact, our naive theories do not often fail to explain the world. This claim is supported by the ontological category formation literature (Keil, 1979; Sommers, 1963). Furthermore, concept formation theorists such as Margolis (1999) suggest that we may mistake a paper bag (nominal kind/artifact) for an animal (natural kind) at a distance but it is soon rectified when given a second look or more time to process the information. We may also need to shift within ontological categories such as reassigning the concept 'whale' from the category of 'fish' to the category of 'mammal'. It is argued that we already possess the category of 'mammal', therefore the shift is a smooth one. The same may even be said of Vosniadou's study of children who eventually shift the concept of Earth from that of a flat object to one that is spherical and rotates around the sun (Vosniadou & Brewer, 1994). Although interesting, these sorts of category shifts are outside the scope of this current study.

The Nature of Instruction

The second challenge involved with this category is identified as the problems of textbook structuring of science content. Chi (2000) identified the following weaknesses: (1) too much emphasis on macro level processes, therefore micro level actions are described in terms of classes of individuals rather than interactions and collective effects; (2) lack of emphasis on how macro level patterns emerge from micro level interactions; (3) insufficient emphasis on emergent processes; finally, (4) inadequate attention and direction concerning when differential strategies are required for problem solving. These problems too are outside the scope of this current study, however, the instructional intervention developed for this current study was sensitive to these concerns and referred to in the discussion section.

2.5.1 Overcoming Lack of Awareness of Emergent Processes

Chi (2000) argues that emergent processes are difficult to pinpoint because they are intertwined with [linear] causal processes. She also suggests that it is difficult because for everyday practical purposes the world is seen as functioning in a [linear] causal fashion. “Thus, the adequacy of operating at the phenomena level in an everyday world, coupled with the absence of any conflicting feedback from such an operation (such as closing the window to keep the heat out), prevents the students from being aware that their interpretation is limited at some deeper level. Without such knowledge and feedback, there is no motivation for students to seek an alternative [nonlinear] explanation” (Chi, 2000b, p.24).

Further support of this conjecture comes from Resnick and Wilensky who have argued, together and independently, that such misconceptions be considered to constitute the “deterministic/centralized mindset” (Resnick, 1994; Resnick & Wilensky, 1995; Wilensky & Resnick, 1999). They posit that attending to a single level of description, rather than the connections among levels, constrains our ability to correctly explain emergent patterns of behavior in both physical, chemical, and human organizational processes (e.g., birds flocking, gas equilibration, traffic jams). They assert that a possible explanation for this predilection is our commanding but myopic human experience as “active planning and designing agents in the world. Yet most of the natural world is composed of agents with much smaller capacities – agents that do not have enough neuronal capacity to conceive of a plan or enough bandwidth to communicate it to conspecifics.” (Wilensky, 2001, p. 3). Adding to this argument, is a common tendency to anthropomorphize these behaviors, perhaps because of our limited experience with these types of agents, or perhaps because of an innate psychological behavior left over from childhood where human traits are used to explain inanimate objects (Vosniadou, 1989).

Yet another possible weakness of human cognition is our inability to reason about multiple operations of very large numbers of entities. Wilensky (2001) states: “Because of our experience as agents and our inability to attend to large numbers of factors or long periods of time, we do not usually have significant opportunities to develop robust intuitions about how emergent phenomena arise and maintain themselves" (p.3).

2.5.2 Summary of Identified Challenges to Learning About Emergent Processes

To summarize, Chi (2000) identifies three inter-related barriers to understanding the ontological category of emergence. First, novice learners treat macro level behavior as the linear sum of causal events (lack of nonlinearity understanding). Second, learners fail to consider the interactions between agents at the micro level (lack of local interactions understanding). Third, learners fail to understand that the macro level emergent behavior is the result of the *collective* interaction of agents *and environment* - "interactions in dynamic collections" (lack of emergent process).

Resnick's and Wilensky's work, together and independently, have echoed these three common limitations of understanding and have added to the list the constraints of the *deterministic* and *centralized* mindset that guides our thinking and reasoning about emergent phenomena (Resnick, 1994; Wilensky & Resnick, 1995; Wilensky, 1993; 1995; 1997; Wilensky & Resnick, 1999). Additionally, recent work by Wilensky's team (Stieff & Wilensky, 2002) supports that still another limitation, the over-attribution of *static properties*, appears to be a contributor to learning about emergent processes.

These studies, along with Jacobson's (2000) expert-novice categorization, support the contention that non-scientific attributions may be articulated into the following six categories of habitual assumptions, and studied as such:

- Isomorphic behavior at both macro and micro levels (i.e., reductive bias or non-awareness of emergence);
- Centralized control assumption;
- Single causal explanation of macro-level behavior from micro-level interactions (i.e., additive, linear);
- Determinacy assumption;
- Intentionality (i.e., teleological, anthropomorphic); and,
- Static outcomes to processes assumption (e.g., beginning-end processes).

2.6 The Status of Research on the Teaching of Emergence and Complexity

At least two major questions rise out of the body of literature described above. One is explicitly theory based and relates to the practical efficacy of Chi's model of conceptual change. The second is related to the instructional strategies that may lead to the development of understanding emergent causation. In essence, two practical questions may be posed: First, can experiential training related to the emergent ontological category facilitate conceptual change (conceptual change defined as a shift in causal explanations)? Second, how else can we come to learn and think about complex systems emergence?

2.6.1 An Empirical Study on Teaching Emergence

So far, the only empirical study which supports Chi's theoretical account of conceptual change was conducted by Slotta and Chi (1999). Slotta examined the effects of a training unit composed of a self-designed computer simulation of dynamic systems related to the diffusion of gases. Using a pretest and posttest experimental design, Slotta randomly assigned 24 university undergraduates with no science background to one of two treatment conditions (experimental and control). There were two sessions of approximately two hours each. The first session presented the ontological training module while the second was intended to provide the science content, transfer material. The training module consisted of computer simulation and text covering four attributes of air expansion and liquid diffusion considered to be an example of "equilibration processes": (1) no clear cause and effect explanations, (2) system of interacting components moving towards equilibrium, (3) combined effects of many smaller processes occuring simulataneously and independently, and (4) no beginning or ending of the process. During the training sessions prompts were used to ensure that students were paying attention to the important parts of the text.

After the training session the subjects were administered both near (air expansion and diffusion questions) and far transfer (predator-prey populations) questions to determine their comprehension of the text and asked to apply the four newly learned attributes to these questions. The test items were muliple choice "problems" on the subject of electric current based on previous work by Slotta et al. (1995).

The control group received the same content on diffusion but no ontological training. The next phase of the experiment provided the students with specially prepared text on electricity based on a conventional physics textbook but with the water analogies removed. The experimental group was cued to transfer by being told: “they would be reading about another example of an equilibration process” (p. 21).

Results from the experimental group showed significant pretest and posttest gains $F(1,22) = 6.8$, $p = 0.02$. In order to tease out the differences between those who understood the material from those who did not, the experimental group was split into high and low scorers on the training posttest. The interaction of test scores with this group was also significant, $F(2, 21) = 13.8$, $p = 0.00$. Further sorting of the data provided better explanations of the findings, which revealed that students’ improvement in problem solving was directly dependent on their understanding of the training.

Slotta also collected verbal protocol data, which was analyzed using the predicate analysis technique developed in Slotta et al. (1995). Novice explanations included the following substance predicates: moves, supplied, qualified, rest, absorbed, consumed. He also identified six expert predicates: system-wide, movement process, uniform state, equilibrium state, simultaneity, and interdependence. The comparison of pretest posttest use of predicates also supported a change due to the training. There was a significant increase in the process predicates $F(1,10) = 31.04$, $p = 0.000$ with a decrease in the substance predicates $F(1,10) = 20.17$, $p = 0.001$. A further refinement in the analysis revealed the same clarification between those who understood the training material and their shift of explanation to the process based ontology. Slotta concluded by stating that:

> Thus, a seemingly tangential training about an ontological category has yielded dramatic results in terms of qualitative reasoning (problem solving and explanations) in another domain (in electricity concepts) that reflects ‘far transfer’ or deep conceptual change (Slotta & Chi, 1999, p. 29).

2.6.2 Using Models to Teach about Emergent Causation

Papert (1980) asserts that our culture is rich in pairs, couples, and one-to-one relationships, however, it is poor in publicized models of systemic procedures. In fact, he states: “Anything is easy if you can assimilate it to your collection of models. If you can’t, anything can be painfully difficult" (Papert, 1980, p. vii). Models, specially computer-based models (e.g., exploratory modeling at described by Wilensky & Resnick, 1999) have proven to be powerful at reifying certain concepts and thereby supporting certain types of learning. Details of three such studies are presented below.

2.6.3 Using Physical Models to Teach About Complex Systems

Hmelo et al. (2000) studied 6th grade students' understanding of the respiratory system using an intervention in which students built partial working models of the lungs. This study shed light on the many *affordances* for promoting deep learning of systems. They adopted a Structure Behavior Function model to both describe the system as well as to code the learners' mental models (based on SBF theory - posited by Goel & Chandrasekaran, 1989). Their results provided evidence that students in the experimental group had a small but significant increase in their attending to structural relationships. Their understanding also became more rich, demonstrated by the number of relationships mentioned and their thinking of how the system worked. However, the students did not mention function as frequently and mentioned behavior least of all. These results were not unexpected, in fact, these researchers anticipated that the causal behaviors would be more difficult to observe because they happen at an invisible level and involved the understanding of dynamic relationships.

Their research is significant because it provides evidence that supports the contention that understanding of complex systems' behavior is possible even by 6th graders. Moreover, it confirms the value of instructional tools with greater affordances for demonstrating emergent processes.

2.6.4 Computer Based Modeling Environments

As previously stated, the research teams of Wilensky and Resnick have conducted many qualitative studies of the affordances realized by different versions of their multi-agent modeling languages (StarLogo, Resnick, 1994; StarLogoT, Wilensky 1997; and the subsequent simulation models of NetLogo, Wilensky, 1999; and, currently ChemLogo). Although the study described below took place after my research was in progress, I present it as an example of one of the few structured inquiries of this particular modeling tool.

Stieff and Wilensky (2002) examined six undergraduate science majors' understanding of the process of chemical equilibrium when using a modeling and simulation package called ChemLogo. This modeling environment is embedded in NetLogo, which may be considered the next generation derivative of StarLogo. The intervention consisted of a three-part 90-minute interview during which the students were asked to first explain their understanding of the Le Chatelier's Principle in chemistry, then during their observations and interactions with the simulation explain the behavior of the molecules that were realizing the behaviors described by this principle; followed by the opportunity to reflect on their reasoning.

Observations reported describe a shift from formulaic problem-solving approaches and rote memorization (which were exhibited in part one of the 90-minute interview), to attempts at conceptual reasoning and justification of answers during and after using ChemLogo. The authors identified four distinct categories of observed changes: (1) defining equilibrium for a chemical system, (2) characterizing factors affecting equilibrium, (3) transitioning between submicro-, micro-, and macro- levels during problem solving, and (4) fluidly moving between various forms of symbolic representation at all three phenomenal levels. The report featured one student's (Andrew) experience and identified a change in his ability to explain and correct his predictions. He was able to deduce correctly how the micro-level events result in phenomena at the macro-level, a change that provided him with greater confidence to deduce other more accurate and reasonable answers.

2.6.5 Emergent Process Models – *Life*[10]

Penner's (2000) study looked at the development of four 6th graders engaged in a nine week after-school instruction investigating emergent properties using the computer simulation *Life* and 'talus slope'. Using a case study approach allowed him to closely investigate the changes in the students' reasoning. He focused on three issues related to how these students come to develop ways of thinking about emergent systems: (1) How did they achieve an understanding of the patterns that develop? (2) Did they recognize that no primary causal factor was necessary? (3) Did they come to distinguish between micro- and macro-levels of descriptions? (4) How did they explain the effects of small changes on the resulting

[10] As a classical example of emergent processes from simple automaton programming, Life could be played both as a paper and pencil game or on the computer.

development? Resnick's (1994) framework related to issues of centralized versus distributed control (i.e., "lead" or "seed") were used to code the verbal protocols.

The importance of Penner's study was the confirmation that a formal taxonomy based on causal mechanisms could be used to code verbal protocols effectively (Penner (2001) used Resnick's conjecture of a "lead" or "seed" attribution). Secondly some of his participants did begin to experience a shift toward emergent causal explanations.

2.7 Main Research Questions Study

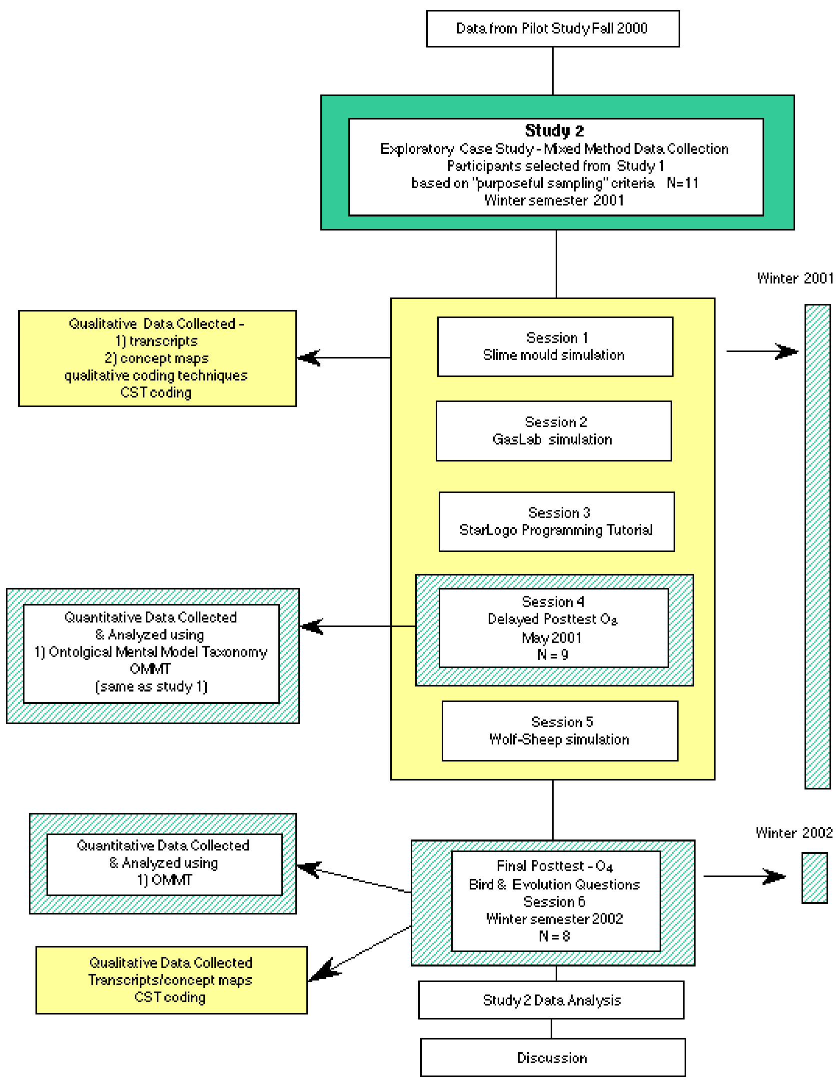
To date, only one known study (Slotta & Chi, 1999) has attempted to answer the first question. While the second question (given below), has drawn attention from a handful of researchers exploring the use of modelling and simulation with varying levels of affordances to learn about emergence (Azevedo, Seibert, Guthrie, Cromley, Wang, & Tron, 2002; Bloom, 2001; Colella, 2001; Hmelo, Holton, & Kolodner, 2000; Klopfer & Colella, 2000; Klopfer & Um, 2002; Penner, 2000; Wilensky & Resnick, 1999; Stieff & Wilensky, 2002; Wilensky, 1999; Wilensky & Reisman, in press; Wilensky & Stroup, 2000). In the study presented here we address three questions:

1. Do student's explanatory frameworks of scientific phenomena (ontological beliefs) change as a consequence of an instructional intervention utilizing simulations of complex systems (i.e., StarLogo) and supported by cognitive scaffolding?
2. What Complex Systems concepts do students acquire during the instructional activities?
3. What is the development of students' systems thinking?

METHODS

3.1 Research Design

The study employed a longitudinal qualitative case study to investigate whether an instructional intervention that employed simulations facilitated the change in students' explanatory frameworks of science (ontological beliefs) from a naïve "clockwork model" to a more expert "emergent systems model". Subsequently, we investigated the influence of the instructional intervention on students' comprehension of specific aspects of complex systems as shown below.



3.1.1 Participants

This study engaged the theory driven sampling strategy of "purposeful" sampling (Creswell 2002). The subjects were selected on the basis of their ability to help the researcher understand conceptual change learning. Motivation and persistence were determined to be an important criterion and therefore subjects with high internal motivation were selected; the assessment of motivation was based on scores from the Learning Approach Questionnaire (LAQ) created by Donn (1989). The participants were nine first year Cegep Science students between the ages of 17 and 18 inclusive.

3.1.2 Instructional Intervention

The experimental intervention consisted of five one-hour sessions that involved the use of different StarLogo simulations: Slime, FreeGas, and Wolf-Sheep. The simulations introduced students to the following component characteristics of complex dynamic systems (as identified by a subject-matter expert):

- The *Slime* simulation introduced the students to: (1) Emergent levels of organization – particularly through visible aggregation and pattern formation; (2) Local interaction of agents; (3) Dynamic homeostatic behaviors and self-organization; (4) Random action of agents; and (5) Small scale fluctuations lead to nonlinear effects.
- The *GasLab - FreeGas* simulation introduced the students to: (1) Emergent levels of organization – however, the macro-level was more abstract and required an understanding of graphs; (2) Local interaction of agents; (3) Dynamic equilibrium and self-organization; (4) Probabilistic nature – particularly related to large numbers and the formation of "normal distributions" (Wilensky & Resnick, 1999); and (5) Small scale fluctuations lead to nonlinear effects.
- The *Wolf-Sheep* simulation introduced students to: (1) Emergent levels of organization; (2) Local interaction of agents; (3) Dynamic equilibrium and self-organization; (4) Flows of resources (i.e., multiplier effect); and (5) Small scale fluctuations lead to nonlinear effects.

Over the period of five one-hour sessions, spanning a 7-week period each of the nine first year Cegep Science students, met individually with the coach and worked with the simulations. Metacognitive prompts were provided as needed. Students were asked to explain their thinking on the characteristics and behavior of both the agents (i.e., the turtles) and the resulting system; a product of the interactions between agents (e.g., slime mold colony, volume & pressure, or ecosystem). The students constructed a set of characteristics that could be identified as common to complex dynamic systems.

The objectives of the coaching were to scaffold the cognitive load of learning the particular aspects and behaviors of these models. Great effort was made to limit any direct instruction unless the participant showed a substantial lack of understanding or frustration (defined as periods of over 10 to 15 minutes without describing or explaining anything new, or taking the discussion in a completely unrelated direction). Therefore students' observation and acquisition of emergent causal processes should be viewed as the outcome of self-directed discovery rather than direct instruction.

3.1.3 Case Study Design

> Case study is a part of scientific method, but its purpose is not limited to the advance of science. Whereas single or a few cases are poor representation of a population of cases and poor grounds for advancing grand generalization, a single case as negative example can establish limits to grand generalization…. Case studies are of value in refining theory and suggesting complexities for further investigation, as well as helping to establish the limits of generalizability" (Stake, 1998, p. 104).

This study employed a longitudinal qualitative case study design described by Merriam (1999) as particularistic, descriptive, and heuristic. Particularistic, because it focuses on a specific instance, event, program or phenomenon; descriptive, because it results in "thick" and rich descriptions of the phenomenon; and, heuristic, because it sheds light on a particular phenomenon thereby leading on toward new meaning and relationships. These are made possible because of data collection techniques that employ methods such as direct observation, interviews, as well as written documents and artifacts produced by the learner (Patton, 1990). Hence, this design allowed us to focus closely on how students reason about the behaviors of a computer driven multi-agent modeling environment portraying different types of complex systems all of which display emergent causal processes. Furthermore, the inductive nature of the design leaned toward theoretical explanations, not limiting itself solely to straightforward descriptions. Finally, multiple cases were used to strengthen, validate and stabilize the findings (Miles and Huberman, 1984).

One of the key limitations of case study design is the sensitivity and integrity of the investigator (Merriam, 1998). Because of limited opportunities for training, and the close proximity of observer and observed, there may be unintentional bias and loss of perspective. Merriam (1998) reminds us that since the researcher is both the primary instrument of data collection and the primary data analysts, attention and accounting for bias is important. Lastly, but not least, in all such research designs ethical considerations must be addressed. Guba and Lincoln (1981, cited in Merriam, 1998) tell us that case writers can make the data say anything they may want. Therefore, both the reader and the authors must be wary of these biases and look for alternative explanations and possible externally imposed agendas; particularly, in policy making, socially and politically driven case study research.

Methods of Data Analysis

According to Merriam (1998), there are three main methods of analyzing qualitative data: (1) descriptive accounting of findings, (2) category constructions, and (3) theorizing; whereas Yin (1994) suggests two general strategies: (1) the descriptive framework, and (2) the development of theoretical propositions. Although using different words, both authors suggest that the descriptive level is the less in-depth analytical technique. At the descriptive level meaning is conveyed through the compression and linking of data, which is then presented in a narrative format. Most case studies generate some type of narrative presentation, however, many strive for the more sophisticated method of analysis involving the construction of categories or themes that captures recurring patterns flowing throughout the data. To emphasize this point, Merriam (1998) states: “category construction *is* data analysis” (p. 180).

Methods of constructing data categories. Categories are not the data themselves; rather they are abstractions derived in both a systematic and intuitive manner. Glaser and Strauss (1967) suggest that the categories should be “emergent” (this meaning should not be confused with the way “emergent” has been used thus far in this study); that is, they should be born out of the data and in so doing be a perfect fit thereby explaining most of the data collected. Categories may also be considered lenses through which the data may be viewed. In many instances, including this current study, categories are informed by the purpose of the study as well as the literature.

Methods of constructing "themes". The next level of data analysis is more abstract and involves the construction of explanations through the linking of categories. In case study research, this is considered the cross-case analysis. Merriam's (1998) description of this process is consistent with the qualitative post-positivist movement. By comparison, Yin's (1994) is reminiscent of the quantitative approaches suggesting the identification of dependent and independent variables. Whichever approach is selected, Yin tells us that "the analysis of case study evidence is one of the least developed and most difficult aspects of doing case studies" (p. 102). This current research viewed this challenge of constructing themes and testing the links between categories as an important part of the data analysis.

Establishing Validity (i.e., Trustworthiness and Authenticity)

Different authors suggest that validity of case studies should be established through a variety of methods (Erickson, 1986; Patton, 1990; Merriam, 1998; Yin, 1994). These include: (1) the collection of different data sources thereby allowing for the cross-validation of findings (i.e., triangulation); (2) the use of two or more evaluators to review material and make independent judgments and interpretations (i.e., inter-rater reliability); (3) an adequate amount of data collected over an adequate amount of time to provide a range of cases (i.e., confirming and disconfirming cases); and (4) accuracy of facts and interpretation of data evaluated by the cases themselves.

3.2 Data Collection, Coding, and Analysis of the Three Outcome Measures

In accordance with Patton's (1990) recommendation, we collected direct observational data (audio and video tapes of the instructional activities) written documents (students' responses at the pretest and posttest), interview data collected after each session, and the students' concept maps at the beginning (week 2) and at the end (week 6) of the intervention. These data were subsequently used to construct three measures, a measure of the students' *Explanatory Frameworks* (Ontological Beliefs), a measure of the students' *Conceptual Understanding of Complex Systems*, and a measure of the *Development of Students' Systems Thinking*. Table 2 describes how the research questions are related to these instruments and to the data analysis.

Table 2. Data Analysis Matrix

Table 2. Main Questions: 1. Do students explanatory frameworks of scientific phenomena (ontological beliefs) change as a consequence of an instructional intervention utilizing simulations of complex systems and cognitive scaffolding? 2. What Complex Systems concepts do students acquire during the instructional activities? What appears to facilitate or restrict the acquisition of these concepts? 3. What is the development of students' systems thinking?			
Sub-Component of Questions	Data	How will these data be analyzed?	What can be said about data analysis?
1.1.a What mental models do students construct of the phenomena under study? Do these mental models change during the course of study? 1.1.b Are there stages or dimensions to mental model development (Vosniadou)? *If mental models mediate what we choose to* ***observe*** *and, in turn, shape our* ***explanatory*** *capabilities (Vosniadou, 1994), then we may look at changes in observations and explanatory capabilities as a way to infer what mental model exist.*	*Longitudinal approach* • transcripts • concept maps Comparative approach • pretest - posttest answers • posttest – expert answers • session 6 – expert concept maps Posttest only approach • evolution question	Evidence of their ability to observe/identify and use the newly introduced concepts. Evidence that their explanations using these new concepts.	As the learner's ability to observe new dimensions of the concept(s) increases, the relationships between the concepts change and therefore their explanations change.
1.1.1 What is "observed" by the student? i.e., what concepts/terms are identified during which simulation. 1.1.2 Do students recognize the same concepts across different simulations/situations or different things from different simulations? *This answer will respond to diSessa & Sherin's notion of invariance and integration*.	*Longitudinal approach* Transcripts from sessions using simulations: Session 1 (Slime) Session 2 (GasLab) Session 3 (Tutorial) Session 5 (Wolf-Sheep) (*within case and cross-case comparisons required*)	1.1.1 Frequency counts - descriptive - using grid of complex dynamic systems' concepts (emergent causal) as identified by Jacobson 2000; Holland 1992; Chi 1999. 1.1.2.a Judged by the way that they are used, which concepts appear to be more deeply understood? By which students? 1.1.2.b What's observed from these concepts and how does it change over time: • surface features	1.1.1 Statement of which concepts are identified and which are not. 1.1.2.a Conclusion drawn about the efficacy of certain simulations in developing certain concepts. 1.1.2.b Conclusion about the level of understanding of the concept based on the number of times used or ways they are

Table 2. Main Questions:
1. Do students explanatory frameworks of scientific phenomena (ontological beliefs) change as a consequence of an instructional intervention utilizing simulations of complex systems and cognitive scaffolding?

2. What Complex Systems concepts do students acquire during the instructional activities? What appears to facilitate or restrict the acquisition of these concepts?

3. What is the development of students' systems thinking?

Sub-Component of Questions	Data	How will these data be analyzed?	What can be said about data analysis?
1.1.3 Are there differences between the students' level of understanding? How is this explained?		• structural features 1.1.3.a Based on the transcripts, who understood which concepts? 1.1.3.b Based on the concept maps who organized concepts more like an expert? Who did not? 1.1.3.c Based on the posttest answers, who used more emergent framework explanations? Who did not?	used. 1.1.3.a Statement of what types of student characteristics may explain the different types of results. 1.1.3.b Statement of what types of understanding may be required to facilitate the learner's change in explanatory frameworks (i.e., acceptance of a new ontological perspective).
1.2 How are the concepts organized?	1.2.1Concept maps from sessions. Number of concepts correctly linked compared to experts. 1.2.2 Transcripts coded for one "focal" concept at a time and see what is observed/discussed with its mention.	1.2.1 Patterns of recognition - are there relational hierarchies, clumps or chains? 1.2.2 Comparison to experts concept maps of concepts.	1.2.1 Statement of which concepts are probably conceptually linked in causal network. 1.2.2 Statement that comparing these learners to experts organization and structuring.
1.3 Is there conceptual change? *Restructuring* (conceptual change) • adding new information • new observations	*Comparative approach* • pretest – posttest • concept maps	1.3.1 Propositional relationships of concepts: i.e., which concepts are related. From a propositional analysis of explanations in each session.	1.3.1 More deliberate use of emergent concepts. E.g., greater understanding of probabilitistic phenomena and the role of uncertainty. Greater understanding of the

<table>
<tr><td colspan="4"><u>Table 2. Main Questions</u>:
1. Do students explanatory frameworks of scientific phenomena (ontological beliefs) change as a consequence of an instructional intervention utilizing simulations of complex systems and cognitive scaffolding?
2. What Complex Systems concepts do students acquire during the instructional activities? What appears to facilitate or restrict the acquisition of these concepts?
3. What is the development of students' systems thinking?</td></tr>
<tr><th>Sub-Component of Questions</th><th>Data</th><th>How will these data be analyzed?</th><th>What can be said about data analysis?</th></tr>
<tr><td>• new points of reference</td><td></td><td>1.3.2 Rate the quality of the conceptual network being developed (high, medium, low).

1.3.3 Pre/post test comparisons of concept use.</td><td>role of interaction among agents in determining organizational patterns of behavior.</td></tr>
<tr><td colspan="4">Emerging Hypothesis A.1: Learners who are able to identify the same concept/terms from different simulations will be more likely to have understood and integrated that concept into their causal network.
(based on diSessa & Sherin's integration and invariance).
Emerging Hypothesis A.2: instruction using StarLogo and cognitive coaching will facilitate the development of new schema of emergent heuristics. This in turn will facilitate conceptual change in this knowledge domain.</td></tr>
<tr><td>1.4 What is retained after a period of no instruction? What changes occur after 8 months?</td><td><u>Comparative approach.</u>
• transcripts 5 compared to 6
• concept maps changes between session 5 and 6</td><td>1.4.1 Recall of concepts and terms. Judgment based on:
• use of concepts in explanations and general answers.
• definitions provided.
• answers to direct questions of recall.

1.4.2 Consistency of terms used in the specific context. Comparison with prior use.
1.4.3Development of concepts within the concept maps.</td><td></td></tr>
<tr><td colspan="4">Emerging Hypothesis B: the level of understanding will have increased due to the embedding of these concepts into other disciplines. Therefore, it becomes more integrated into declarative knowledge networks and becomes Context Independent knowledge (Barsalou, 1987).</td></tr>
</table>

3.2.1 Explanatory Framework (Ontological Beliefs) Taxonomies

After the students listened to an introductory lecture on complex systems they were asked to give a written response to a problem involving a complex system (the pretest). The students' written responses were coded as described below and used to determine the changes in their explanatory frameworks of scientific phenomena (ontological beliefs). Approximately 5 weeks after the pretest, the students were given a second, similar problem involving complex systems (the posttest). The problems are described in Appendix A.

An ontologically-based coding rubric with a reasonably high reliability ($\alpha = 0.81$) was developed based on Chi, Slotta & deLeeuw (1994) and Jacobson (2000). It was used to code the students' responses on the pretest and posttest. The specific taxonomies used to code the *clockwork mental models* (CWMMs) and *emergent framework mental models* (EFMMs) are presented in Tables 3 and 4.

Table 3. Specific taxonomy used to code Clockwork Mental Models (CWMMs).

Clockwork Mental Model (CWMM)	**Components of coding**
Ontological perspective – *Reductive*	1) Agents' act in isolation. 2) Simple stepwise description.
Control of system – *Centralized*	• Orders/controls come from outside. Or is within the system but not attributed to the individual agents within. (e.g., different agents have different rules; mention of hierarchy).
Action effects – *Linear*	• One thing leads to another. E.g. direct link between controller and controllee. (e.g., action→reaction)
Agents' actions – *Deterministic (i.e., Predictable)*	• Agents' actions are predictable. e.g., they (it) will perform the action. There is no mention of randomness or chance in their actions.
Underlying causes – *Teleologic*	• It knows the end point: e.g., it knows it has to survive.
Systems' Nature – *Static*	• Explicit descriptions of non-changing system.

Table 4. Specific taxonomy used to code Emergent Framework Mental Models.

Emergent Framework Mental Models (EFMM)	Components of coding
Ontological perspective – *Emergent Self-organization ontology* *Question: 1.Does a pattern emerging? 2. Is there a difference between agents and system? 3. What draws the system together?*	1) Local interactions among agents, 2) leads to the creation of something that exhibits a differential behavior than those of the component agents; 3) this interaction is made possible due to some type of identification (tagging device /organizing agent), 4) and, communication (flows of information and/or resources).
Control of system -*Decentralized control* *Question: Who or what initiates the formation of the system?*	1) The individual agents are independent of each other, yet they all operate under the same rules;
Action effects – *Nonlinear effects* *Question: Are there feedback loops within the system? Do they amplify or control the outcome?*	1) Positive feedback is a feature of these systems therefore small actions can exhibit exponential results.
Agents' actions – *Random action (indeterminacy)* *Question: How do the agents behave before they are part of the system?*	1) Agents appear to act in random independent fashion, *Also possibly present in the answer:* 2) Randomness allows for variability and variety within the system.
Underlying causes – *Probabilistic causes (Stochastic)* *Question: Is the same outcome guaranteed each time the system forms?*	1) The system organizes itself based on the interactions of the agents as described above, therefore the resulting structure is never certain, rather it is stochastic which implies that there is a probability based emergent pattern. *Also possibly present in the answer:* 2) Like other probabilistic processes, larger numbers over longer time periods are more likely to result in the formation of normal distributions.
Systems' Nature – *Dynamic homeostatic nature* *Question: Is there movement of the agents within the system?*	Agents may move through, and in and out, of the system, however the system persists in a self-organizing fashion. 1) Once the system, the recurring structure, emerges it exhibits a more stable quality; yet all the component agents have the potential to be replaced by other similar independently operating agents.

The raters first read the entire response made by each student and parsed it using the following procedure adapted from Mosenthal and Kintsch (1992a, 1992b): identify the verb (action), the noun (agent), and the object (agent effect). Using the mental model taxonomy, the raters next coded each parsed statement into one of three possible categories (EFMM, CWMM, or NM) according to how it answered the following five questions: (1) who or what is controlling the system – "Control of System"; (2) how do the agents' behave at the start of the process – "Action Effect"; (3) what are the effects of the agent's actions – "Agents' Actions"; (4) what is the underlying cause of the system's behavior – "Underlying Cause"; and (5) how does the system behave – "Systems' Nature"? (see Tables 3 and 4).

The raters coded the student's responses into the sixth category, "Ontological Perspective", somewhat differently. Because of the more global nature of this category, it was the only one to be coded at a large grain level. That is, raters coded the entire response rather than not each parsed statement. The strategy for scoring this category relied on the answers to the following four questions: Was there mention of (a) local interactions between agents, (b) identification of some mechanism that would draw the agents together (tags), (c) recognition of a flow of information or resources which creates the system out of independent agents (flows), and, (d) identification of some type of pattern formation as the agents come together to form a system? If there were answers to any of these four questions the appropriate letter was entered. If, however, the student's response did not address these questions, but instead there was evidence of a stepwise (reductive) approach to the explanation, coupled with descriptions of the agents as isolated entities (i.e., no interaction among agents) the answer was coded as CWMM. As before, if neither mental mode applied then the NM column was coded.

From the initial testing of the coding scales it was determined that it was easier to code each question twice: once to identify evidence of one mental model (e.g., CWMM), then again to identify evidence of the other mental model (e.g., EFMM). This method produced greater consistency from the raters. To clarify by way of example, rater #1 started coding all pretest responses for evidence of emergent framework mental models (EFMM). He would then repeat the process a second time, coding all pretest answers for evidence of clockwork mental models (CWMM). Rater #2 would by contrast, start coding all pretest responses for evidence of CWMMs, then repeating the process, code for evidence of EFMMs. This method addressed possible threats of an "order effect" from the coding procedure.

All the data derived from the ontologically-based mental model coding were scored according to a binary method (1 or 0) thereby indicating evidence (1) or no evidence (0) of a particular mental model. Although this binary method causes a loss of sensitivity[11] due to the inability to distinguish the relative frequency of idea units (Slotta et al., 1995), its use was justified because of the increase in clarity when it came to identifying the learner's likely mental model (i.e., EFMM, CWMM, or NM).

The question of reliability was addressed by having three raters (rater #1 = a biology graduate student, the primary rater for the final coding, rater #2 = the biology subject mater expert, and rater #3 = the first author). All raters coded the posttest responses, however, only raters #1 and #3 coded the pretest responses.

The training of the raters took approximately 60 minutes and they were provided with a coding key (see Appendix B). The inter-rater reliability was established by comparing the individual coding of two raters. The number of total responses was multiplied by the number of categories to be coded, then by the number of possible mental model stated (EFMM, CWMM, NM). Differences between raters were counted as the raw number of cells that were different. Therefore, if one coded a category as EFMM while the other coded the same category as CWMM this was counted as two changes.

On the pretest scores there was agreement on 418 out of 450 scores yielding an inter-rater reliability coefficient of 0.93. On the posttest scores there was agreement on 140 out of 162 scores, yielding an inter-rater reliability of 0.86. Inconsistencies were resolved through discussion between raters until consensus was reached.

3.2.2 Conceptual Understanding of Complex Systems

A coding schema was also developed to determine students' conceptual understanding of the various aspects of complex systems (i.e., the Complex Systems Taxonomy referred to as the CST). It was adapted from Jacobson (2000) and from the study of complexity (i.e., complex systems) and reflects concepts presented by Holland (1995, 1998), Bar-Yam (1997), Kauffman, (1995), and others. It was intended to provide a "Fine Grain" description of the behaviors that emergent phenomena exhibit; therefore, most of the overlapping of concepts was removed. The CST was used to code the transcripts because it provided the broadest list of categories that could be identified from these data. It was used to examine the student's focus of attention on aspects of complex systems and as such, tells us more about how their understanding of complex

[11] Slotta, et al. (1995) in reference to the loss of sensitivity state: "this measure, although still yielding the same basic results, robbed the analysis of any sensitivity to differences between or patterns within the use of different predicates" (p. 386).

systems may have occurred. In essence, it sheds light on: (1) which concepts may be easily understood, in the process helping to explain the results of this as well as other similar studies; (2) which concepts may be strongly entrenched in component beliefs; and (3) which may not be well represented by the intervention. The complete CST is presented in Table 5. Further elaboration on the process of how the coding schemas were developed appears in Appendix C.

Table 5. Complex Systems Taxonomy (CST).

COMPLEX SYSTEM CODING TAXONOMY	DEFINITIONS
1. Local interactions.	Local interaction of many individual agents result in the formation of higher level entity.
2. Simple rules produce complex results.	Simple rules produce complex results through a complex interplay of local interactions.
3. Decentralized control	Emergent systems exhibit organization without centralized control. (i.e., decentralized control). Agents are independent and competition among themselves.
4. Random behavior	Behavior of the individual agent is uncertain because of the innumerable possible local interactions.
5. Tags	Tags are an organizing mechanism that allows the agents to select among agents or objects that would otherwise be indistinguishable. They are filtering, specialization, and cooperation devices. Tags are the mechanism behind hierarchical organization-the agent/meta-agent/meta-meta-agent/...organization.
6. Flows • Feedback (positive or negative) .	Flows of information/resources throughout the system using networks involving: node → connectors → resources. Flows through networks vary over time. Moreover, nodes and connections can appear and disappear as the agents adapt or fail to adapt. Feedback is where the influence of an element impacts on other elements through a series of relationships that return to the initial point, i.e., feeds back on itself. *Multiplier* effect is the result of (positive feedback). Example of positive feedback (amplification of the initial state. Out of control if it goes to far). Helps achieve contained contraction or replication and growth or can lead to uncontained and unstable contraction or growth. *Recycling* effect - also defines the constraints.
7. Internal models	Internal models (schemas) give the agent the power to anticipate - tacit internal models simply prescribes a current action/ overt internal models uses lookahead protocols.
8. Diversity/ variability	Diversity also known as "requisite variety". A control system must have adequate variety. The variety of the control system must be greater than the variety of the controlled system or the environment.
9. Modularity	Hierarchical nature of systems allow for recycling of useful components. Building blocks are the components of a complex system that can be used and reused in a great variety of combinations like a set of Lego building blocks. These reusable components make it possible to make sense of novel situations. Subassemblies are building blocks of the emergent complex system.
10. Pattern formation	Pattern formation: Prominent among simple mathematical models that capture pattern formation are local activation / long range inhibition models.
11. Open/closed systems	Generally emergent systems are open systems but can be closed (e.g. gas pressure).

COMPLEX SYSTEM CODING TAXONOMY	DEFINITIONS
12. Multiple Levels	Systems are nested. Therefore complex systems are made up of many subcomponents that may themselves be nested systems.
13. Probabilistic	Probabilistic behaviors have non-deterministic outcomes. Population size affects the results. The larger the sample size the more reliable the prediction of outcome and the more the outcome reflects a "normal distribution" curve. The smaller the sample, the more likely that individual differences will make it difficult to predict the outcome.
14. Nonlinearity	Nonlinear systems are more complex than linear systems. A feature of nonlinear systems is that different starting points lead to different end points and can cause the model to become unstable. Behavior is often counterintuitive.
15. Criticality	Lever points wherein small amounts of input produce large directed change; threshold effect (e.g., phase changes).
16. Dynamic equilibrium	Homeostasis /Dynamic equilibrium with fluxes in and out. The notion that organisms (systems) exchange information, materials and/or energy with their environment in order to survive therefore over time the materials that make up the organism (system) has partially or totally changed. Multiple (meta) stable states. Small displacements (perturbations) lead to recovery, larger ones can lead to radical changes of properties. Dynamics on such a landscape do not average simply. Mathematical models are generally based upon local frustration e.g.. spin glasses, random Boolean nets. Attractor networks use local minima as memories. Examples: weather - persistent structures, proteins - results of displacements in sequence or physical space, physiology - the effect of shocks, dynamics of e.g. the heart, brain/mind - memory, recovery from damage, economy/society - e.g. suggested by dynamics of market responses.
17. Adaptation	Adaptation is defined as agent and environment interactions. An example is "fitness landscape". "Fitness landscape": Part of a Hill-climbing algorithm in which the search space turns into a fitness landscape, where every point in the space ("horizontal") is associated with a "vertical" fitness value, so that a landscape with valleys and peaks appears. Problem-solving then reduces to "hill-climbing": following the path through the fitness landscape that leads most directly upward.
18. Selection	Selection suitability of the particular trait an agent has for surviving long enough to reproduce in a particular environment. It is also defined as information (a la Shannon's theory). Selection as information is relevant to the issue of multiple selection: replication (reproduction) with variation, and comparative selection (competition) as a mechanism for POSSIBLE increase in complexity. Consistent with modern biological views of evolution it is essential to emphasize that selection does not have to increase complexity.
19. Time scale.	Time scale is a critical feature in development of complex systems.
20. Multiple causality	Emergent systems are dependent on multiple actions and interactions to create their complexity. Therefore the number of agents in an environment with multiple components to interact with will create infinite possibilities of outcome.

3.2.3 Conceptual Development of Students' Systems' Thinking (Concept Maps)

Adhering to the three criteria – a task, a response format, and a scoring system – described by Ruiz-Primo and Shavelson (1996) helped to document the concept mapping process and thereby created an audit trail for this event.

The task. Ruiz-Primo and Shavelson, (1996) tell us that the "task" is defined in three parts, (1) what the learner is required to accomplish, (2) the constraints of the performance, and (3) the "task content structures". The latter is described as "the intersection of the task demands and the constraints with the structure of the subject domain to be mapped" (p.578). In this study students' task demands were straightforward: construct a concept map reflecting their understanding of complex systems' behaviors. The task constraints included using the terms provided, however, it was not limited to those terms only. The students were allowed to arrange the terms in any manner that best reflected their changing understanding. Lastly, no constraints were placed on the structure of the maps therefore the final organizational structures were evaluated as evidence of students' conceptual understanding.

Response format. In this study the student was presented with the 12 terms related to complex systems' behaviors on "post-it" notes and asked to arrange them on a board in such a way as to express their understanding of how the terms may be related. Maps from the first mapping session (session 2) therefore appeared a little clumsy because of this technique, even though students were given the opportunity to draw links or add comments. After the initial activity, the interviewer transcribed the maps into pencil and paper representations. All subsequent mapping activities were made in this dual mode with the student provided first with the paper version of their map, and if they required more freedom to move nodes around, they were allowed to use the post-it notes on a board. These two modes of response formats were viewed as supporting each other, therefore, they should not account for any variation or change in the concept maps produced between students or sessions.

Criteria Used to Score Concept Maps

Informed by the literature (Ruiz-Primo & Shavelson, 1996; Jonassen, Reeves, Hong, Harvey, & Peters, 1997), we selected two criteria to evaluate the students' concept maps: (1)

concept pairings (McClure & Bell, 1990), and (2) organizational structure (DeSimone & Schmid, in press).

Criterion 1. The importance of the concept-parings (i.e., node-link-node relationships) was established using criterion maps (experts' maps). The literature tells us that comparisons to experts are always controversial because there is substantial evidence that experts' knowledge representations vary dramatically one from another. However, we attempted to address this limitation by using an averaging technique (e.g., average of experts, average of high achieving students, etc.) as described by Acton, Johnson and Goldsmith (1994). Their findings suggest that average ratings of experts improved the comparisons. Because of our access to a limited number of experts we used a standard textbook definition of complex systems as a starting reference point.

Establishing the weighting of the links based on the text definition. Among the twelve concepts, three were lexical concept: "complex system", "simple system" and "system". Of the possible pairings it was anticipated that "complex systems" lexical concept would form a central node and be closely linked with the following concepts: decentralized, dynamic, random, self-organizing and probabilistic (represented as "yellow" rectangular nodes in the concept maps); whereas, "simple systems" would not be directly linked to these terms (represented as "blue" rectangular nodes in the concept maps). Therefore the former relationships were assigned a score of 1 point for each direct link. The remaining concepts, centralized, algorithmic, static, predictable, (represented as "white" rectangular nodes in the concept maps) could be linked to either "simple systems" directly, or to "complex systems" but in an indirect fashion; that is, qualified by direction of links and/or propositional statement between nodes. No score therefore were assigned to these and less predictable paired relationships.

Establishing the weighting of the links based on criterion maps. Based on the maps collected from four experts, we determined the weightings to assign for the links between nodes. Starting with the map (Figure 1) we established that many concepts were indirectly linked to the central node "complex systems". Furthermore, in addition to that node, the term "self-organization" also formed a central node on other experts' concept maps (see another example

Figure 2). Three of the four experts generated maps supporting these linked pairs. Hence, these consistently paired relationships were assigned a weighting of 3 points. They are as follows:

self-organizing	paired with	probabilistic / random	= 3 pts
self-organizing	″	dynamic	= 3 pts
probabilistic	″	random	= 3 pts

Two levels of pairings were thus assigned. The former as described by the textbook definition of complex systems, and the latter based on these averaged criterion-map paired relationships. If these terms were linked to a paired concept in more than one fashion, each link was scored as a separate pairing.

<u>Criterion 2.</u> As a second criterion to examining and scoring the students' concept maps we drew upon DeSimone and Schmid's technique for analyzing the deep structure and quality of concept maps (see DeSimone & Schmid, in press). However, because concept mapping in this study was used primarily for eliciting verbal protocols (i.e., more elaborated conceptual reasoning thereby richer transcript data), as opposed to being used as a main assessment activity, we chose to adopt a simplified version of their analysis and scoring technique. Instead of examining the maps at the many possible levels of labeled relationships structures, we applied the scoring only to the organizational appearance of the map. Hence, hierarchical maps were assigned 3 points, cluster formations were assigned 2 points, and chain formations assigned 1 point. In the event of maps that were somewhere between a cluster formation and a hierarchical formation, we assigned a score of 2 points. Only obvious vertical relationships with evidence of subsuming levels of organization were assigned as hierarchical maps.

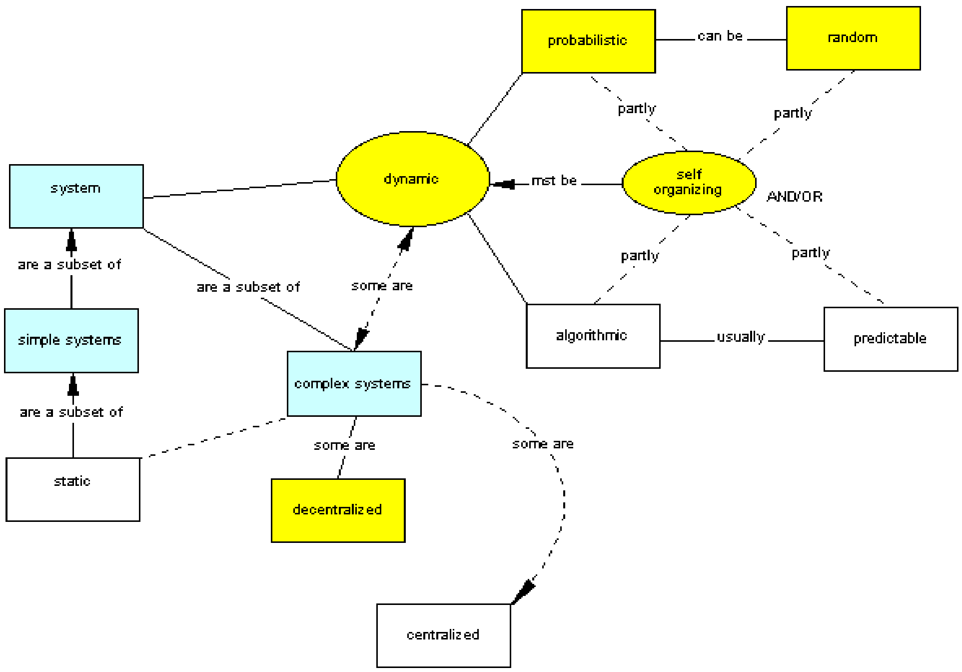


Figure 1. Expert's concept map of complex systems concepts.

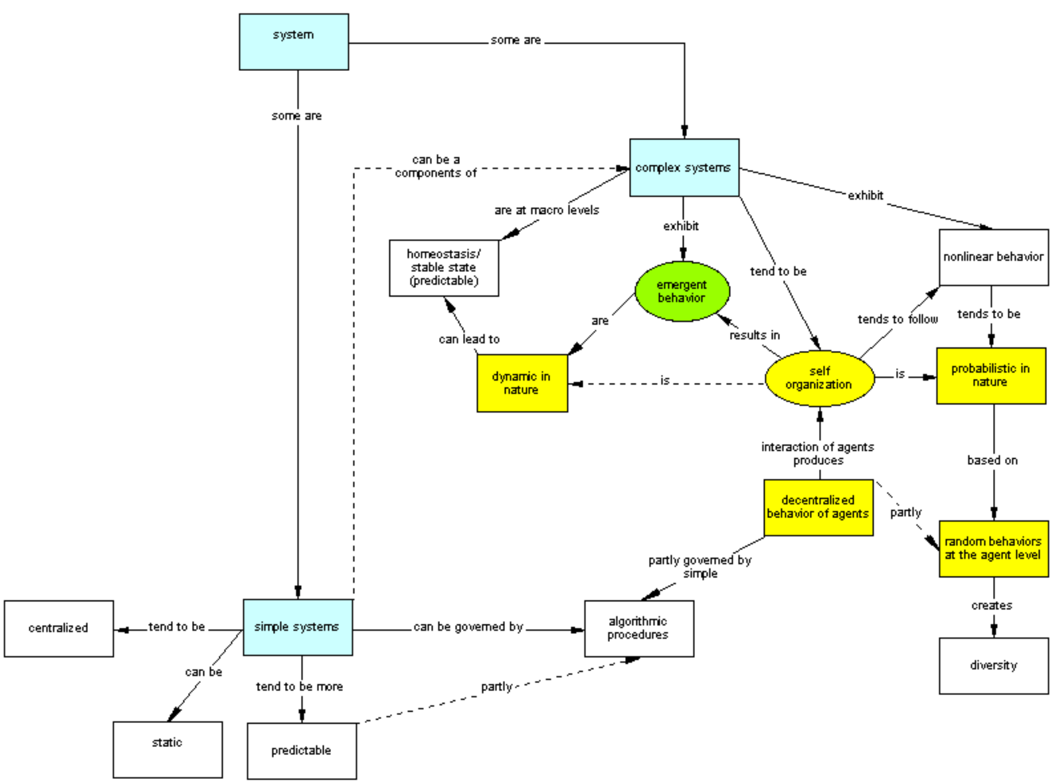


Figure 2. Advance-learner concept maps of complex systems concepts.

RESULTS

We addressed the following three questions in the research described in this paper:

1) Do students explanatory frameworks of scientific phenomena (ontological beliefs) change as a consequence of an instructional intervention utilizing simulations of complex systems (i.e., StarLogo) and supported by cognitive scaffolding?
2) What conceptual understanding of complex systems do they acquire during the instructional activities?
3) What is the development in students' systems thinking as evidenced by their concept maps?

4.1 Question 1: Changes in Explanatory Frameworks

Table 6 and Figure 3 show that there was an increase in students' responses indicating emergent framework mental models (EFMMs), but, little or no change in students' responses indicating clockwork mental models (CWMMs). At the same time there was a significant decrease in the number of responses that could not be coded as either mental model.

Thus, ontological training supported the acquisition and use of emergent framework mental models (EFMMs). In fact, all nine students gained some level of awareness of emergent causal processes and all demonstrated some level of transfer ability in solving emergent analogous problems. Clockwork mental models (CWMMs) do not change substantially as a result of emergent ontology instruction, and possibly emergent frameworks mental models indeed do not replace clockwork mental models. Students with no mental models of emergent processes may benefit more from ontology training than students with clockwork mental models.

Table 6. Gains in students' mental models (n = 9).

Scores on	Pretest		Posttest	
	M	SD	M	SD
Emergent Framework Mental Models (EFMMs)	5.6	5.6	12.7*	5.8
Clockwork Mental Model (CWMMs)	3.4	3	3	2.8
No Model (NM)	10.7	3.1	5.9*	3.2

* Significance at α 0.01 on one-tailed independent samples *t* test.

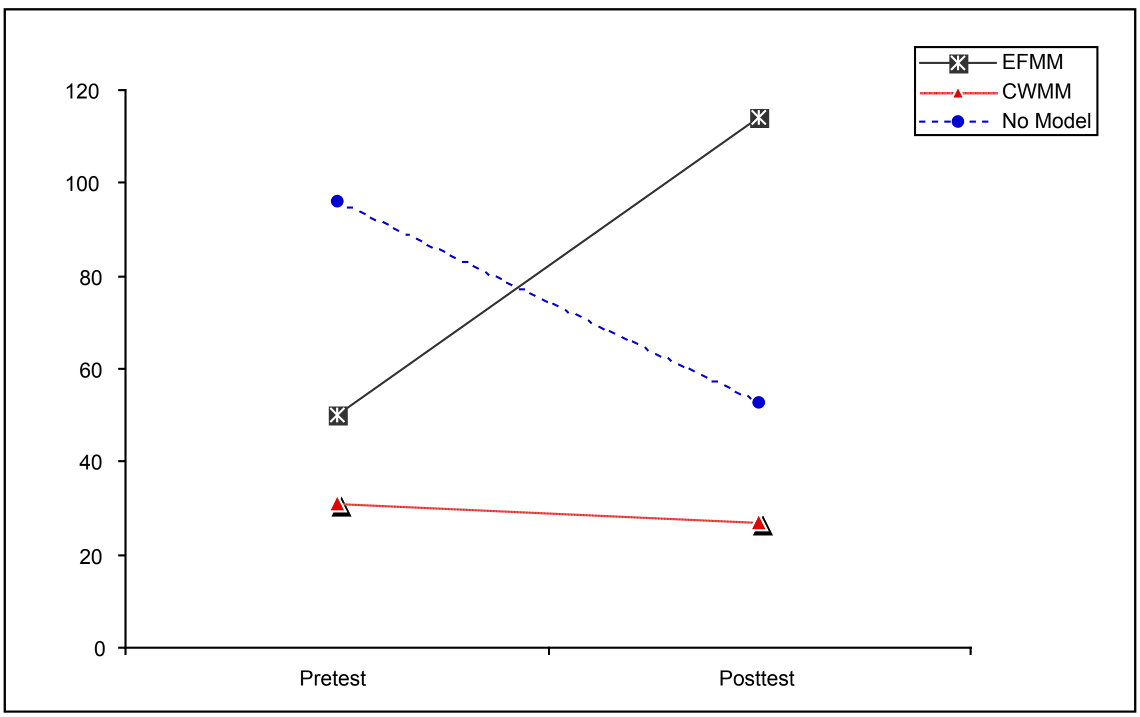


Figure 3. Changes in Explanatory Frameworks or Ontological Beliefs

Table 7 shows a categorization of the changes in each student's mental model and indicates that most students showed evidence of synthetic mental models. This supports Vosniadou et al.'s (2001) contention that this may be a necessary part of the conceptual change process.

Table 7. Meaning of changes to mental model categories.

Mental Model			Status	Label	Students
Emergent Frame-work (EFMM)	Clockwork (CWMM)	No Model coded (NM)			
+	–	–	Desired state. Learning of EFMM occurs. It is stable and coherent (i.e., integrated into understanding).	Emergent Mental Model Group	Greg
+	+/–	–	With gains in both mental models, they are unstable, but the learner is aware of their explanations therefore more coherent.	Synthetic 1A Mental Model (increasing coherence)	Mitch Sam Sidney
+	–	+/–	With losses in CWMMs the learner is moving towards a more stable mental model but with gains in NMs it means that they are unsure and therefore lack coherent understanding. Therefore they cannot bring their ideas together to generate coherent explanations.	Synthetic 2A Mental Model (increasing stability)	Walter
+/–	–	+/–	With gains and losses in EFMMs, and gains in NMs it suggest that the mental models are unstable and incoherent	Novice Emergent Mental Model	Emilie Norman Penny Monique

(increase = +, decrease = –)

4.2 Question 2: Conceptual Understanding of Complex Systems

The results in Table 8 represent the scores for 13 out of the possible 20 categories of the complex systems taxonomy (CST) for all nine students over the course of three instructional sessions that employed simulations. From this we can make the following two generalizations:

- When we collapse across students, and focus on the total scores for each session, we can see that the specific simulation used influenced which complex systems concepts were discussed by the students. We call this dimension the affordance of the simulation. Thus, differences among the simulations influence the effectiveness of the intervention.
- When we collapse across simulations, and focus on the combined relative scores, we can see that although all the students discussed, and possibly learned, some of the complex systems concepts, there was variation in both the nature and number of complex systems concepts discussed by the students. Thus, individual differences among students influence the effectiveness of the intervention.

Table 8. Complex systems concepts identified by student and reported by session.

Simulation	Student	Complex System Concept													Combined Scores
		ML	LI	OS	PR	RB	FL	TA	DE	SR	DC	DI	NL	PA	
Slime	Norman	8.4	4.6		2.3	2.1	1.3	0.4		0.1					**19.1**
	Penny	9.1	2.2	0.7	2.2	0.7	1.1			0.1	1.5			0.4	**17.9**
	Emilie	12	2.4	0.3	0.1		0.4		0.1						**15.4**
	Monique	6.7	2		0.3		0.4	0.2	0.3	0.1	0.1				**10.1**
	Walter	9.6	4.4	0.5	2.3	0.3	1.1		0.3		0.2				**18.9**
	Mitch	12.5	6.6	1.1	3.2	0.6	1.5		0.2						**25.7**
	Sidney	12.2	6.3	0.3	4.1	2.5	0.4	0.4	0.1	0.3	0.2	0.3			**26.8**
	Greg	11.4	7.8	1.3	5.2	1.6	1.3	1	0.3	0.4		0.3		0.7	**30.9**
	Sam	11.5	6.4	1.1	2.1	1.9	0.5			0.4	0.5			1.6	**25.6**
Total scores		**93.4**	**42.7**	**5.3**	**21.8**	**9.7**	**8**	**2**	**1.3**	**1.4**	**2.5**	**0.6**	**0**	**2.7**	**190.4**
GasLab	Norman	3.7	2.7	0.7	1.5	0.7					0.2				**9.5**
	Penny	4.4	2.1	0.6	1.7	0.1			0.2		0.1		0.1		**9.3**
	Emilie	4.1	1.4	0.3	0.9	0.2									**6.9**
	Monique		1.0		2.0				0.2						**3.1**
	Walter	4.8	2.3	1.4	2.0				0.1				0.1		**10.6**
	Mitch	5.8	5.6	3.7	3.0	1.3		0.2	0.2		0.2		0.2		**20.2**
	Sidney	4.3	3.2	1.4	4.3	0.4			0.2		0.1			0.1	**14**
	Greg	8.4	7.9	6.8	5.1	0.2		0.2	0.2				0.2	0.2	**29.2**
	Sam	5.6	3.1	1.5	2.0	0.6	0.3	0.1	0.3		0.2				**14**
Total scores		**41.1**	**29.3**	**16.4**	**22.5**	**3.5**	**0.3**	**0.5**	**1.4**	**0**	**0.8**	**0**	**0.6**	**0.3**	**116.8**
Wolf-Sheep	Norman	4.0	8.0	2.0	2.0	0.3	0.3	0.5	0.1				0.1	0.5	17.8
	Penny	2.2	2.3	0.8	1.3	0.1		0.2		0.1	0.3				7.5
	Emilie	4.9	4.7	1.6	1.6	0.3		0.1			0.6				13.8
	Monique	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Walter	6.5	6.3	2.2	2.4	0.7		0.7	0.2		0.1		0.1	0.8	**20.2**
	Mitch	7.7	8.4	0.6	5.4	0.8		1	0.2		0.1			0.1	**24.4**
	Sidney	14.3	7.6	0.7	2.4	0.8		0.5	0.1	0.1	0.2	0.1	0.1		**27.1**
	Greg	6.0	10	4.0	2.7	0.6		0.8	0.5		0.3	0.1	0.3		**25.7**
	Sam	5.2	8.0	1.5	3.2			0.8			0.6	0.4		0.2	**20**
Total scores		**50.8**	**55.3**	**13.4**	**21**	**3.6**	**0.3**	**4.6**	**1.1**	**0.2**	**2.2**	**0.6**	**0.6**	**1.6**	
Grand totals		**185.3**	**127.3**	**35.1**	**65.3**	**16.8**	**8.6**	**7.1**	**3.8**	**1.6**	**5.5**	**1.2**	**1.2**	**4.6**	**307.2**

ML is Multiple Levels of Organization, **LI** is Local Interactions, **OS** is Open Systems, **PR** is Probabilistic Behavior, **RB** is Random Behavior, **TA** is Tags, **FL** is Flows, **DE** is Dynamic Equilibrium, **SR** is Simple Rules, **DC** is Decentralized Control, **DI** is Diversity, **NL** is Nonlinear, **PA** is Pattern Recognition

Affordances of three simulations. One of the major themes constructed from the categories to emerge from the interviews was that the different simulations facilitated the acquisition of different aspects of complex systems. That is, the simulations differed in their affordances for learning emergent causal processes.

The results in Table 9 represent the total responses aggregated across students. It displays the percentage of responses within each complex systems component.

Table 9. Distribution of Responses within Complex Systems Taxonomy (CST) for each simulation.

Simulations	Percentage of responses within each Complex Systems Concept												
	ML	**LI**	**OS**	**PR**	**RB**	**TA**	**FL**	**DE**	**SR**	**DC**	**DI**	**NL**	**PA**
Slime	49.1	22.4	2. 8	11.4	5.1	4.20	1.10	0.68	0.74	1.30	0.32	0.00	1.40
FreeGas	35.2	25.1	14.0	19.3	3.0	0.26	0.43	1.20	0.00	0.68	0.00	0.51	0.26
Wolf-Sheep	32.5	35.3	8.6	13.4	2.3	0.19	2.90	0.70	0.13	1.40	0.38	0.38	1.00

ML is Multiple Levels of Organization, **LI** is Local Interactions, **OS** is Open Systems, **PR** is Probabilistic Behavior, **RB** is Random Behavior, **TA** is Tags, **FL** is Flows, **DE** is Dynamic Equilibrium, **SR** is Simple Rules, **DC** is Decentralized Control, **DI** is Diversity, **NL** is Nonlinear, **PA** is Pattern Recognition

We constructed the following 4 point scale to describe the affordances based on the experts' ratings, the literature, and the assumption that all the simulations should have offered equal affordances for learning the six major aspects of complex systems described by Jacobson (2000).

- *High affordance* = above 34%
- *High Moderate* = 33% - 17%
- *Low Moderate* = 16% - 8%
- *Low* = 7% - 3%

Table 9 indicates that all three simulations provided affordances of at least high moderate for two complex systems concepts: "multiple levels of organization and "local interaction". They provided a low moderate affordance for the complex system concept "probabilistic nature" and a low affordance for the complex systems concept of "random behavior". The other categories in the coding taxonomy produced fewer than 3% of the total observations and therefore were considered to indicate a weak affordance for learning.

All the simulations provided opportunities to observe and discuss "multiple levels of organization", "local interactions" and "probabilistic nature" albeit to different extents. Only the Slime simulation provided opportunities to observe and discuss "tags" but it did not provide opportunities to observe and discuss "open systems"; the Wolf-Sheep simulation did not provide opportunities to observe and discuss "random behavior". All the simulations had only weak affordances (< 3%) for all the other complex systems concepts. However, despite the similarities in the affordances of the simulations, students did not acquire the complex systems concepts equally well from all simulations as will be discussed shortly.

4.2.1 Individual Difference among Students

Figure 4 illustrates the combined scores on the CST for each student across all sessions. On this basis students could be classified into four groups:

- *Sophisticated Emergent Causal Processes (ECP) Identifier* (CST score > 75). This describes Greg who is considered an outlier at the high end.
- *High Moderate Emergent Causal Processes (ECP) Identifier* (CST score between 60 and 70). This describes Mitch, Sidney and Sam.
- *Moderate Emergent Causal Processes (ECP) Identifier* (CST score between 40 and 50). This describes Walter and Norman.
- *Novice Emergent Causal Processes (ECP) Identifier* (CST score between 30 and 40). This describes Emilie, Penny, and Monique (an outlier at the low end).

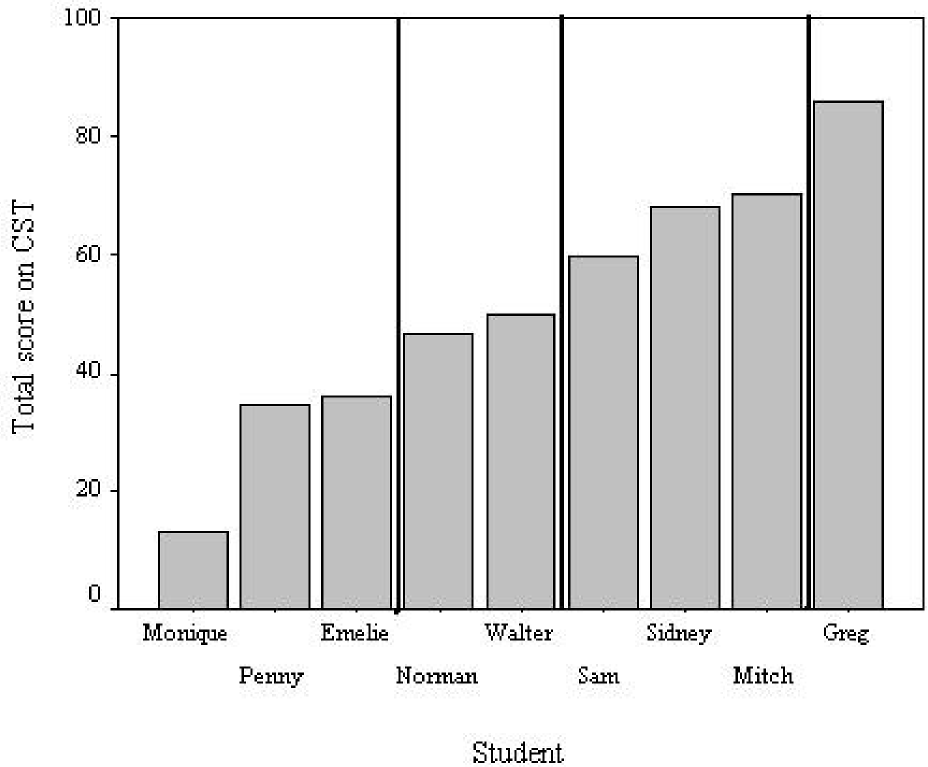


Figure 4. Student's understanding of Complex Systems concepts over three simulations

The results of Table 10 show the number of statements (relative to each student's total number of statements) that were coded (using the CST) into each Complex Systems concept. Thus, it allows us to make a provisional decision on whether each student observed and therefore discussed the Complex Systems concepts. If one arbitrarily, takes a value of 1 as the cutoff point, we can provisionally conclude that all students including the three Novice ECP Identifiers (Monique, Emilie, and Penny) observed and discussed the concepts of "multiple levels of organization", "local interactions", and "probabilistic causes". All the other students also observed and discussed the concept of "random behavior". The major difference between the Moderate ECP Identifiers (Norman and Walter) and the High ECP Identifiers (Sam, Sidney, and Mitch) was in the general strength of their responses. On the other hand, the Sophisticated ECP Identifier (Greg) not only had a greater response to the latter concepts, he also observed and discussed more concepts, namely "flows" and "dynamic equilibrium".

Table 10. Relative number of statements made by each student coded into Complex Systems concepts over the three simulations.

	Novice ECP			Moderate ECP		High-Moderate ECP			Sophistic.
Concept	Monique	Emelie	Penny	Norman	Walter	Sam	Sidney	Mitch	Greg
ML	6.7	21.0	15.7	16.1	20.9	22.3	30.8	26.0	25.8
LI	3.0	8.5	8.8	11.3	13.0	17.5	17.1	20.6	25.7
OS	0.0	2.2	3.6	2.7	4.1	4.1	2.4	5.4	12.1
PR	2.6	2.3	3.6	5.8	6.7	7.3	10.8	11.6	13.0
RB	0.0	0.5	0.9	3.1	1.0	2.5	3.7	2.7	2.4
TA	0.4	0.4	1.1	1.6	1.1	0.8	0.4	1.5	1.3
FL	0.2	0.0	0.0	0.4	0.0	0.1	0.4	0.2	1.2
DE	0.5	0.1	0.2	0.1	0.6	0.3	0.4	0.6	1.0
SR	0.1	0.0	0.2	0.1	0.0	0.4	0.4	0.0	0.4
DC	0.1	0.6	1.9	0.2	0.3	1.3	0.5	0.3	0.3
DI	0.0	0.0	0.0	0.0	0.0	0.4	0.4	0.0	0.4
NL	0.0	0.0	0.1	0.1	0.2	0.0	0.1	0.2	0.5
PA	0.0	0.0	0.4	0.0	0.8	1.8	0.1	0.1	0.9

ML is Multiple Levels of Organization, **LI** is Local Interactions, **OS** is Open Systems, **PR** is Probabilistic Behavior, **RB** is Random Behavior, **TA** is Tags, **FL** is Flows, **DE** is Dynamic Equilibrium, **SR** is Simple Rules, **DC** is Decentralized Control, **DI** is Diversity, **NL** is Nonlinear, **PA** is Pattern Recognition

Interpreting the results of the low counts for categories that exhibited low moderate (random behavior) or weak affordances (nonlinear effect, decentralized control, dynamic equilibrium, etc.), as indicating that students do not observe these concepts may not necessarily be the correct conclusion to draw. It may be argued that a low level of observation data does not necessarily mean a lack of awareness or understanding of a concept. Indeed, it may indicate that the learner readily recognized the behavior described by the concept and chooses to focus instead on other concepts that were more challenging or interesting. It may also mean that the simulation does not lend itself to observing and therefore discussing the concept.

A further qualitative investigation of the students' experiences helps us to understand more about what variables are important to the learning of emergent causal processes. In other words, it allows us to closely examine which of the variables (e.g., the intervention, the learner, the conceptual knowledge) played a greater role in the acquisition of a conceptual understanding of the various aspects of Complex Systems. Thus, we will describe the acquisition of the two concepts with at least high moderate affordances across the three simulations. They are "multiple levels of organization", and "local interactions".

4.2.2 Acquisition of Concept of Multiple Levels of Organization

The multiple levels of organization within the system are an important component of the aggregation emergence process. A typical observation for this category would be that individual agents behaved in one manner while systems (meta-agents) behaved in a different manner. Observation of differential behaviors often resulted in terms such as "random" or "unpredictable" individual agent actions versus "stable" or "orderly" meta-agent actions. Additionally, descriptions could extend to homeostatic behaviors, although that concept was a unique node on the complex systems coding taxonomy (CST).

Therefore what the students appeared to learn was that individual turtles behave randomly (Slime simulation) or unpredictably (other simulations) but when they form a meta-agent or change states, this new level of organization behaves in a more stable and predictable fashion. This observation created some constraints in the students' understanding of where the term predictable fit into a concept map related to complex systems. This topic will be expanded upon in the discussion.

It should also be noted that the data analysis for the FreeGas and Wolf-Sheep simulations was not as clear-cut as for the Slime simulation. In essence, coding for levels of organization required reading several phrases to capture the distinction that the students made between either the agent and the meta-agent attribute of net changes (e.g., pressure, total energy within the system), as described for FreeGas; or between the individual and population levels as well as the level of the system itself, as described in Wolf-Sheep. For instances, key terms or phrases such as "population", "system", and references to the graphical readout "parabolic curve", "sine wave", to list a few, were identified as evidence of reference to the macro-level.

All three simulations provided opportunities for the students to observe and discuss multiple levels of organization as seen in the student interviews presented below. However, there was a decrease in the students' discussion this concept over the three simulations (see Figure 5).

Of the nine students, eight appeared to have understood this concept. Only Monique appeared to have not understood multiple levels of organization. Below are some of the students' statements from each simulation.

Greg (Slime Simulation): yeah, you can see that sometimes they leave but the whole system stays constant no matter if they leave or join the clusters the system seems to follow the same pattern. 273

Emilie (Slime Simulation): The group would be functioning kind of, well, at this, at this point the group would be functioning kind of, you know, in a stable way, but the individuals like some of them would not be...would not be...would not...I don't know how to put this.
647

Sam (FreeGas): Cause there's only a certain amount of energy in the entire thing. It's just that the individual molecules changes... All together [however] they own the same amount of energy throughout. 426

.

Sidney (Wolf-Sheep Simulation): I don't think it was the simpler one with which we had the altruistic and the selfish ones, because in that one it did not have to eat grass, or uh eating something. And in that case when, where you added more of one, the other one didn't counter-balance it, so there's not kind of an equilibrium state. And what's really odd is the population is somehow, the entire population, or like the population of the... sheep and the wolves has reached some kind of constant stability. They're going up and down but in a small, a small frame. And I would call this like uh, a stable state. Because... 177

I: Mm-hm. 179

Sidney: ...because what's happening is at, at one point the sheep are...the sheep population is greater than the wolf population, and the next moment it's the reverse. 181

(pause 5s)

Sidney: But this is where I have everything equal, where I have the, the initial sheep and the initial wolves the same. But I want to try let's see now. Now I'm going to make the wolves' population twice the number.

I: OK, but what do you think is going to happen once you do that? 187

Sidney: At first I would think perhaps maybe now that the wolves would, would finally overpower the sheep, but somehow I don't think so. 189

(pause 4s)

Sidney: You can see what's happening once again. They've reached, they're reaching the, these are gonna, like a parabolic curve they're taking. 195

Sidney: Well that beats the top one. And once again they've reached some kind of equilibrium state. 199

Figure 5 indicates that there is a decrease in students' observation and discussion of multiple levels of organization. The Slime simulation may provide more perceptual cues and more tightly coupled interactions than the other simulations. Moreover, the Wolf-Sheep simulation presented a relatively greater difficulty in observing the abstract meta-agent levels.

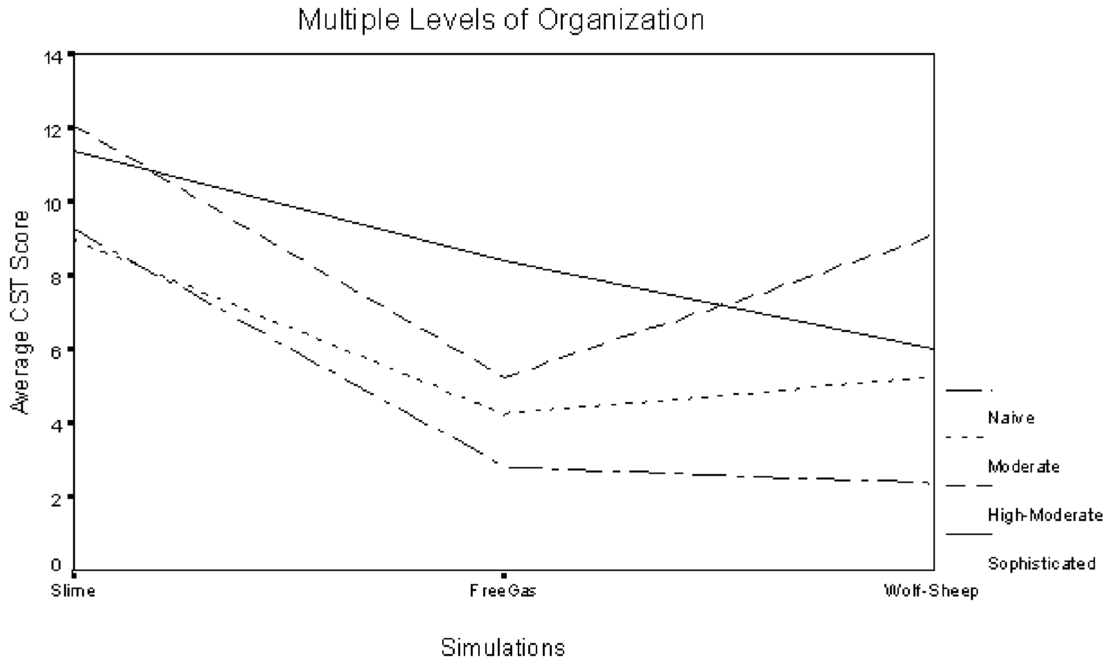


Figure 5. Changes in the number of statements on multiple levels of organization.

4.2.3 Acquisition of Concept of Local Interactions

"Local interactions" refers to the behaviors of agents as they operate within the environments of the simulations. In order to be coded at this node, the student had to generate evidence of their awareness of individual agents (e.g., slime mould, molecules, sheep, etc.), and that these agents' actions were not isolated from each other but in fact significant changes arose because of the different ways in which they could encounter each other or affect each other. In short, the student needed to demonstrate that they could reason about both the individual behaviors of the agents as well as the impact of potential interactions such as attractions, or collisions. In order to code for this category we adopted several key predicate indicators many of which appeared to be used regardless of the simulation content: *interact, attract, collide, aggregate, come together, hit across, form, react to, cluster, move towards, affect, communicate with, organize, change each other*. There were also another group of predicates used to demonstrate local interactions, which suggest telltale signs of "teleological beliefs" that constrained the development of the concept "probabilistic causes": *find, look for, build, and join*.

All the students, with the possible exception of Monique, appeared to understand the concept of Local Interactions. Some examples of these are provided below:

Walter (**Slime** simulation): well, there are a bunch of turtles moving around randomly and they seem to be giving off these green secretions in their wake and it seems as when they get a strong enough secretion they kind of like come together. They kind of like attract each other. 5

Penny (**FreeGas** simulation):: Yeah. If you look at just look at the molecule itself, it's consistent because, since if, well, if they're hitting each other then they're affecting which event takes place. 1049

Mitch (**FreeGas** simulation):Because, it is a system that... Whereas, a collision between a fast and a slow particle, it... They're all elastic collisions, so the energy is considered, in one way or the other. So every elastic collision's going to change the speed of each particle in the system, but it's not going to change the speed of the whole system. So, that's why the average speed is staying around 8.8 or 8.9... 11

Walter(**Wolf-Sheep** simulation):: well, the sheep eat the grass. The wolves eat the sheep. 116

Walter: obviously, when there's no sheep there's going to be no more wolves eventually because they're all going to die out. Because they're not going to have nothing to eat. 169

Figure 5 and the students' statements indicate that each simulation provided a different opportunity to learn about the concept of local interaction. Awareness and possible learning observed from the slime simulation may be attributed to perceptual level display of tightly coupled interactions and aggregation of mould into colonies. On the other hand, local interactions in the wolf-sheep simulation may have been observed because of the expected causal change of events, "sheep eat the grass. The wolves eat the sheep". Hence, one explanation for the increase in observations may be because students are intuitively familiar with these moderately coupled interactions that make up ecological systems.

Another explanation, however, may be a consequence of their changing ability to observe (i.e., "readout" strategies) different causal processes. In this interpretation of the data, the increased number of observations would be described as a consequence of better understanding of relationships between agents and not necessarily attributed to the representational affordance of the simulations. Alternatively, it may be an interaction of the two. Both alternative explanations need to be explored further through other data in this current study and perhaps in future studies.

Figure 6 also indicates that the FreeGas simulation offered weaker affordances of local interactions. These results suggest that dissipative complexity models (whose organization is most apparent at statistical means) may be less likely to generate observations of local interaction, which is congruent with the type of complex system represented by the model. Consequently, it should be viewed as one step towards establishing credibility of the findings. In fact, the modest level of local interaction recognition may be explained by the cognitive scaffolding that cued students to look for similarities across simulations; therefore, in no way attributable to affordances to promote the concept of local interaction generated by this simulation.

Although the FreeGas simulation had weaker affordances for local interactions, Mitch and Greg, nonetheless showed high percentage of their time focused on this concept (43% and 61%, respectively). Both students belong to a cohort of high academic achievers and both had above average scores in NYA physics and chemistry: Greg, 97 (intro. chem. 99), Mitch, 84 (intro. chem. 86). Sidney also achieved a high grade in chemistry (80) but is reported to be aware of this for a mere 23% of his observations. However, his grades in physics (70) suggest that there may be an interaction between the student's level of understanding of physics (specifically collisions) and the ability to observe local interactions. Further investigation into this possible relationship is required.

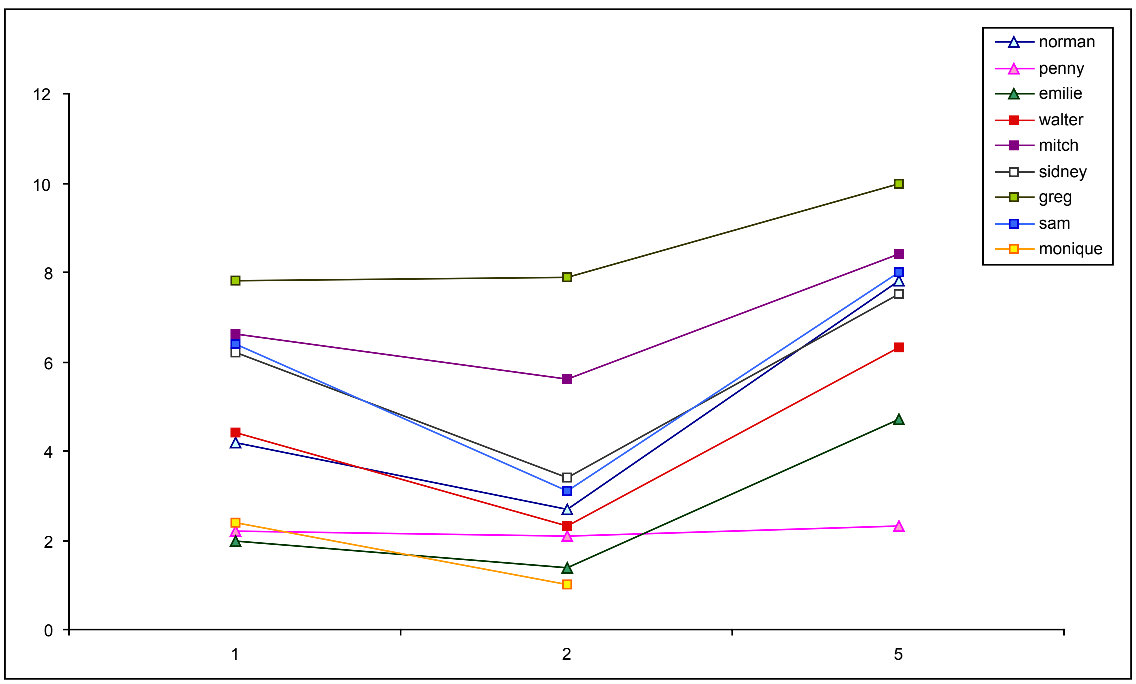


Figure 6. Changes in the number of statements made by individual students on local interactions.

4.3 Question 3: Development of Students' Systems Thinking

The students' concept maps were scored and the results are presented in Table 11. Examining the scores for session 6 (post intervention) the results show that Greg scored 12 points (Figure 7) and Mitch scored 11 points (Figure 8) and made a considerable number of desirable concept-pairings; followed by Sam with 9 points (Figure 9) and Walter with 9 points (Figure 10). By comparison, Sidney scored 7 points (Figure 11) and Norman scored 5 points (Figure 12) both displayed a moderate degree of desirable concept-pairings. Finally, Emilie with 3 points (Figure 13) and Penny with 2 points (Figure 14) both scores are substantially lower than the average. Monique's maps are not shown because she did not complete session 5 and 6, which were crucial in the development of the other students' maps.

Table 11. Scoring of Concept maps on criteria 1 & 2 over time.

Quantitative Scoring of Concept Maps		Norma	Penny	Emilie	Walter	Mitch	Sidney	Greg	Sam
Session 2 - initial	Criteria 1	1	1	0	2	3	4	2	2
	Criteria 2	1	1	1	2	2	2	2	2
	Sub total	2	2	2	4	5	6	4	4
Session 6 -post	Criteria 1	3	1	0	7	9	5	9	7
	Criteria 2	2	1	3	2	2	2	3	2
	;ub total	5	2	3	9	11	7	12	9
Total		7	4	5	13	16	13	16	13
Change		+3	0	+1	+5	+6	+1	+8	+5

In summary, the change over time resulted in large gains for Greg, and moderately large gains for Mitch, Sam, and Walter. Whereas, Norman's results showed moderately small gains, Sidney and Emilie showed very small gains. Penny was the only student to show no change between assessments.

Although these quantitative results allow for comparison across cases as well as triangulation with other results, additional and rich information was gained from the following qualitative inspection of the maps. The concept maps created by the students in session 2 and the final meeting, session 6 are presented on the following pages. These maps almost speak for themselves in as much as they show some dramatic changes in understanding of emergent causal concepts that are framed within the context of complex systems thinking. As the scores above suggest, the maps clustered the students into three groupings: (1) "*sophisticated*", (2) "*moderate*", and (3) "*novice*", understandings of complex system relationships. The maps are presented in decreasing order of elaboration therefore the first ones with represent the most sophisticated understandings.

This quantitative scoring method established a baseline of difference within cases (time series), as well as allowed for between students comparison. Afterwards, the concept maps were examined in a qualitative fashion for evidence of qualitative changes both across time, as well as cross-case analysis.

4.3.1 Group 1, "Sophisticated" Understanding of Complex Systems

Both Greg (Figure 7) and Mitch (Figure 8) produced hierarchical type maps and demonstrated elaborated understanding of the term "complex systems" and its relationship with the concept of "self-organization". In particular, both students appeared to recognize the important connection between it as a central node and the many other associated influences; for example, random action (Mitch) and probabilistic linked to random behaviors (Greg). A further level of understanding was revealed by Greg's addition of the term "emergence", also viewed as a central node. He independently chose to add this term, and, as we can see, he connected it directly to "complex systems" as well as "self-organization". I contend that, from the perspective of this study, this was an important conceptual shift.

<u>An important consideration</u>. Although not accounted for on the scoring schema, how students came to understand this relationship between "emergence" and the other complex systems concepts merely through observations and interactions with the StarLogo environments was important to this study. The intentional omission of this term therefore provided a means to assess the sophistication of the students understanding. (N.B., when the term is present in the student's concept map it is identified as a "dark grey" (or green if in colour) elliptical shaped node).

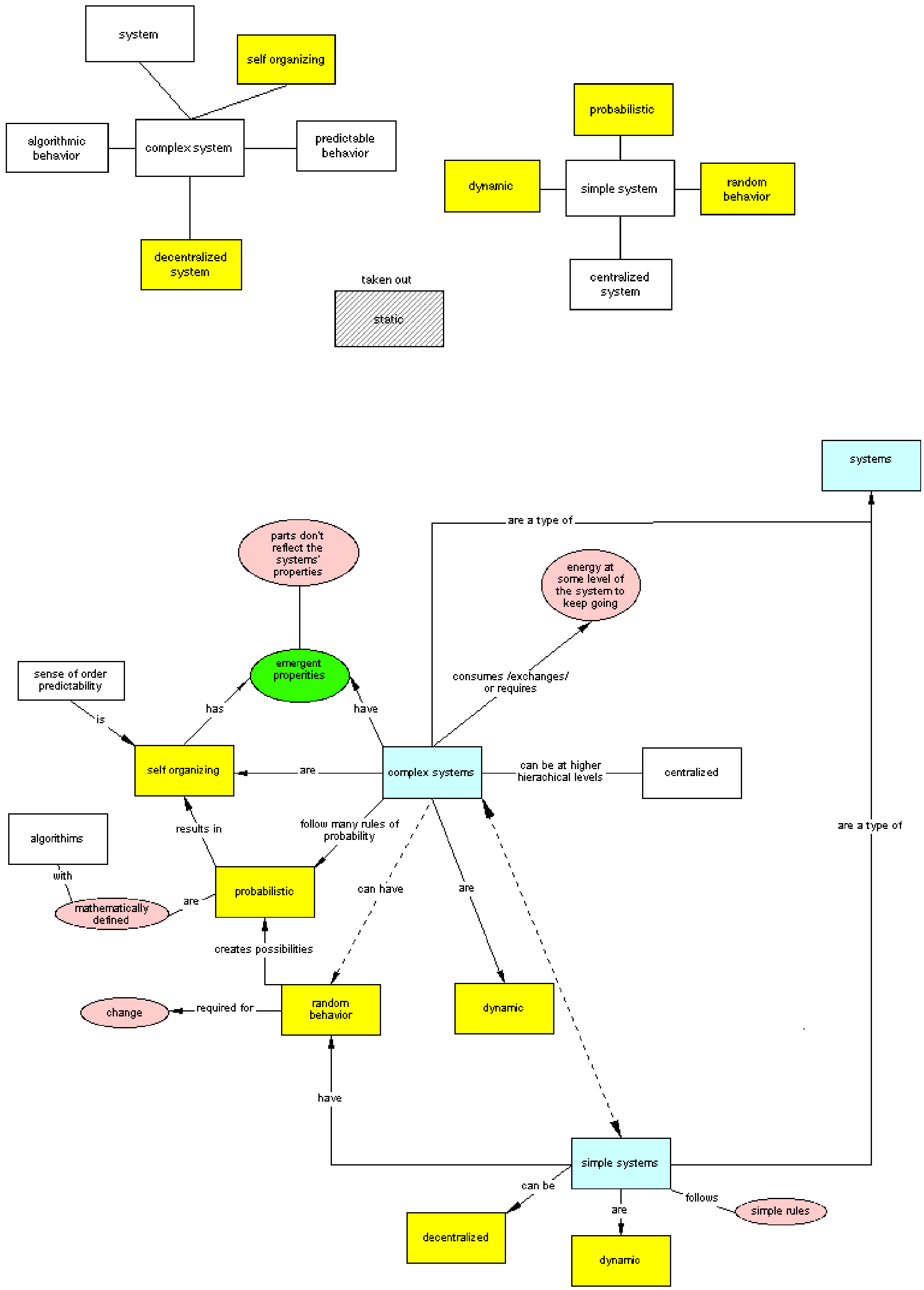


Figure 7. Greg's concept map - Session 2 & 6 (session 2 upper, session 6 lower).

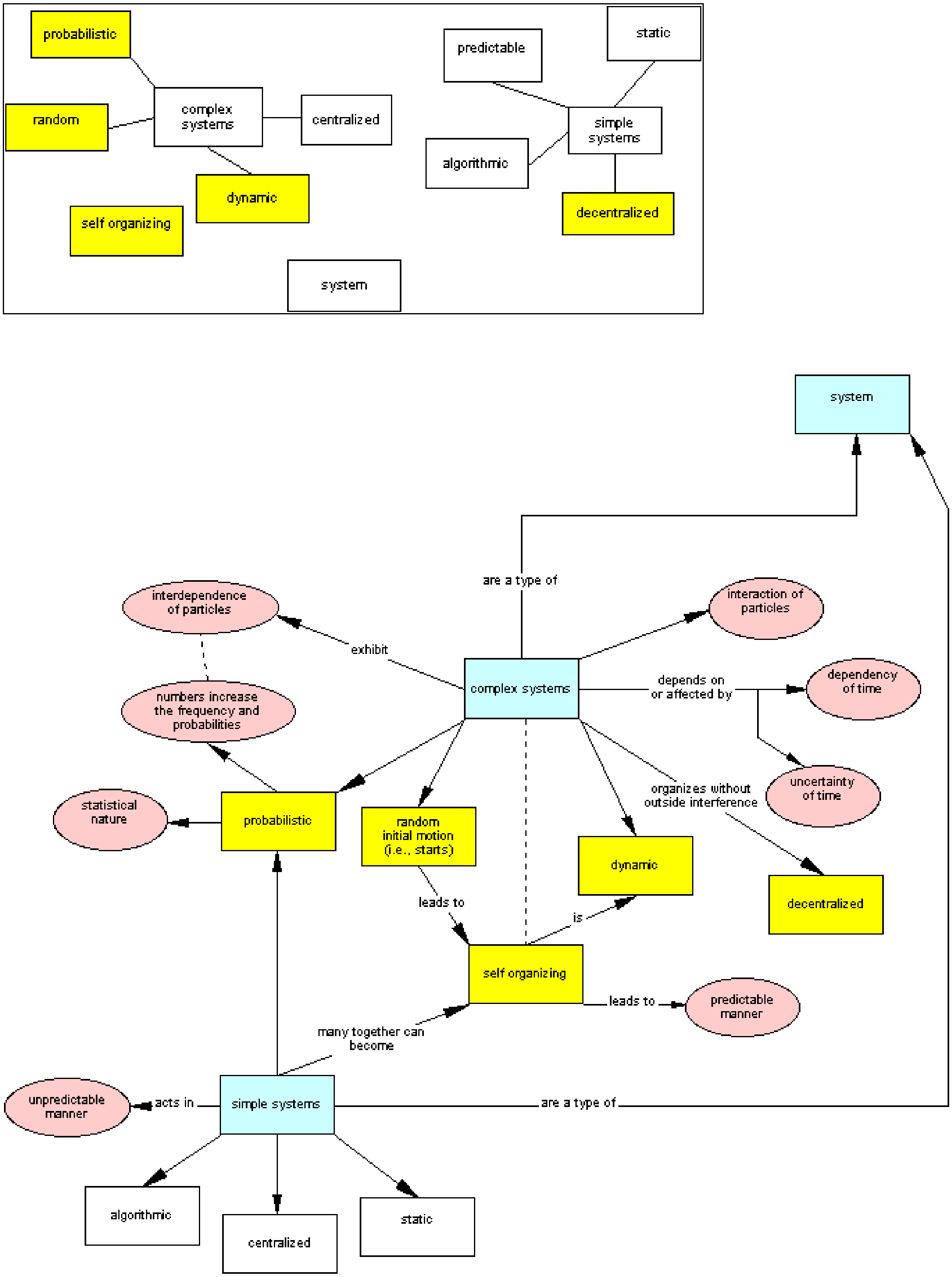


Figure 8. Mitch's concept map - Session 2 & 6 (session 2 upper, session 6 lower).

Sam's concept map (Figure 9) was a cluster formation and looked very different from those described above. However, his organization of terms revealed a level of sophistication that placed him in a unique sub-category of group 1. His identification of concepts was more similar to the expert (Figure 1) in that he created a central node out of both "self-organization" as well as "dynamic". Because of these relationships Sam's clumping of terms into equilibrium-type complex systems and dynamic-type complex-systems demonstrated a higher-level view of complexity. When asked to explain this organization, Sam was able to be reflective and in fact made a small revision to his map based on the conversation below. This is also observable from Sam's score on the CST (Table 10).

I: What's happened between then, which is a long time ago and now (laughs) to make you change? 199

Sam: I think I changed a lot of the way I studied. So I tried to link things a bit together. Not like a concept map, but I tried to link... I do try to go from cause to effect, the results a lot, I guess. So I use just complex system as the basic centre, and I'm working off of that. So I decided that dynamic and static were two different kinds of complex systems, and that all these others were descriptions of a dynamic system while, centralized and all these others were descriptors of static system. And, there, I don't, I don't agree with what I said at all anymore, because to me a static system doesn't have to be a simple system. 201

Sam: Systems that are self-organizing don't have to change that much, and if it's predictable, it's static, if it's algorithmic behavior it follows rules, so it's got to stay pretty much, in a certain area, it's not dynamic like a system that might, rely on chance and random behavior, and isn't centralized.

I: OK, and why did you take away self-organizing? You don't think a dynamic system can be self-organizing? 207

Sam: I think it can... On a bigger scale, I guess. If you look at it from a big scale, a system that relies on random behavior, might, organize itself in the end, but I mean... 209

(pause 4s)

Sam: You see I hadn't thought of that. Now I'm not sure anymore. 213

I: No, no, I'm not suggesting that you're wrong... 215

Sam: I know, but you brought it up and...(laughs) I didn't question that.218

Sam: Because a dynamic system you look at on a bigger scale, to be self-organizing. While a static system I think is more self-organizing, but on a smaller scale. 221

Sam: Pretty much. Like I'm saying it goes, up and down. Like if you're looking at a chart, and it goes up and down, it's not continuous, if you're looking at it from far enough away, it does look like, one straight line I guess, if you're looking at it from far enough away, and that's... That's the way I was thinking about it, I guess. 229

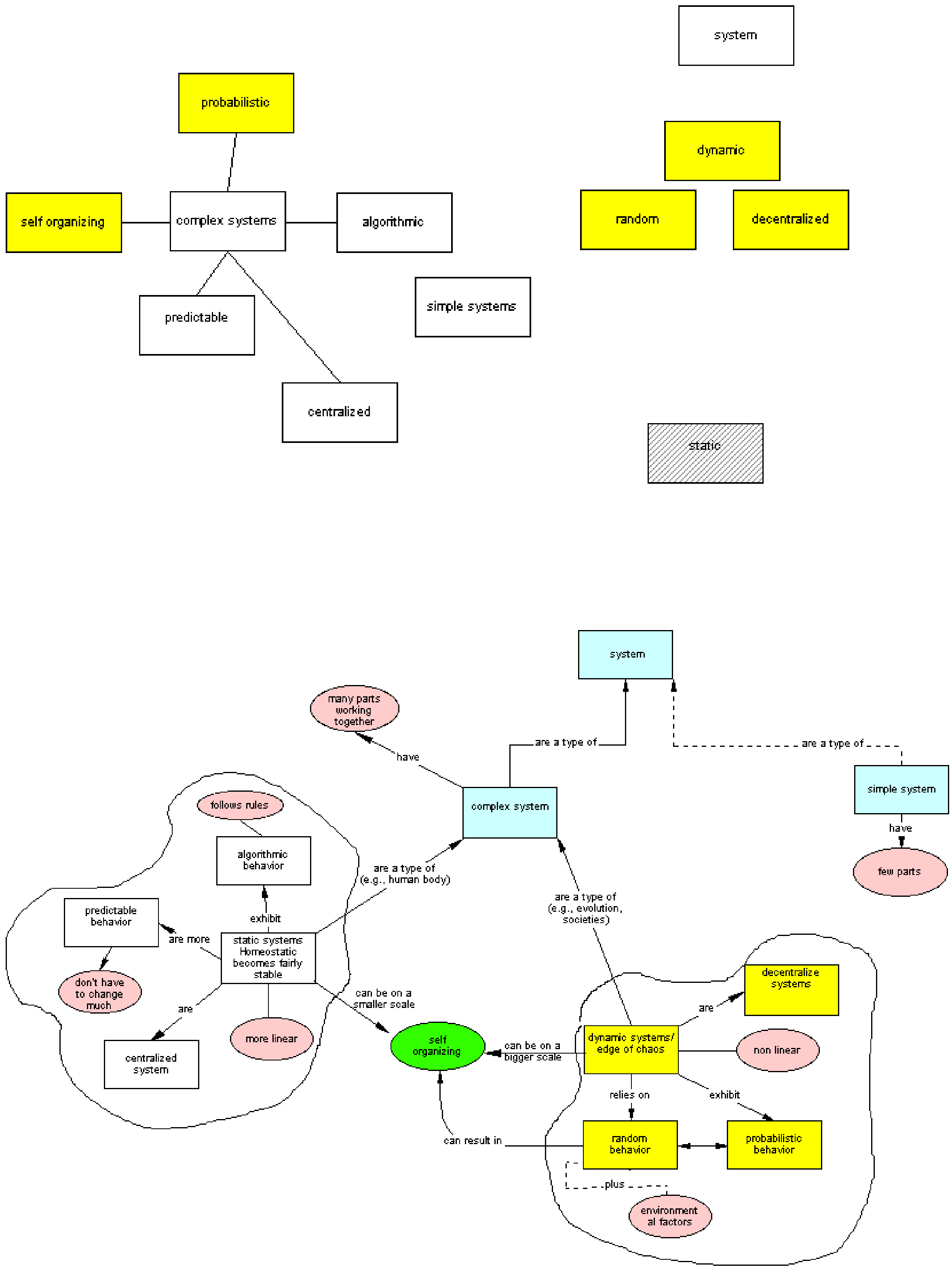


Figure 9. Sam's concept map - Session 2 & 6 (session 2 upper, session 6 lower).

4.3.2 Group 2, "Moderate" Understanding of Complex Systems

Walter, Sidney and Norman all fell into this classification. Whereas Walter's (Figure 10) map is more similar to Greg and Mitch (see above) in its somewhat hierarchical organization, Sidney and Norman produced a cluster type concept map. Walter like Sidney, views "self-organization" only as a link between types of systems – simple system and complex systems. Compared to Sidney (Figure 11), Walter demonstrated some understanding of the probabilistic nature of complex systems and the role of random actions. As argued in the section above on scoring, this was an important recognition therefore placing Walter's understanding above that demonstrated in Sidney's maps

Sidney does not have many important linked relationships. For instances he did not suggest any link between "self-organization" and "random" or "probability". In fact, most of his node-link relationships were single attachments radiating out from the central nodes of "complex system" and "simple system". Nonetheless, he demonstrated that he had elaborated his understanding demonstrated through the additional terms attached (shown in pink).

Finally, both Walter and Sidney added the term "emergent" to their maps, suggesting that they were aware of this concept as a central feature of the intervention. In fact, Walter described why he added the term in the excerpt below.

I: Yeah, it wasn't part of the set of things I gave you. The ones in grey were the ones that you added. Which I thought was a good thing to add. 237

Walter: Yeah. 239

I: I was very pleased that you added those. 241

Walter: Well yeah, because we observed these things in the computer programs that we were looking at. 243

I: And what's an emergent property again, how would you describe that? 245

Walter: From what I remember it's just, um from observing the system there's certain properties that are characteristic of that system, that you begin to notice after you observe it for awhile. 247

I: OK and an example...?

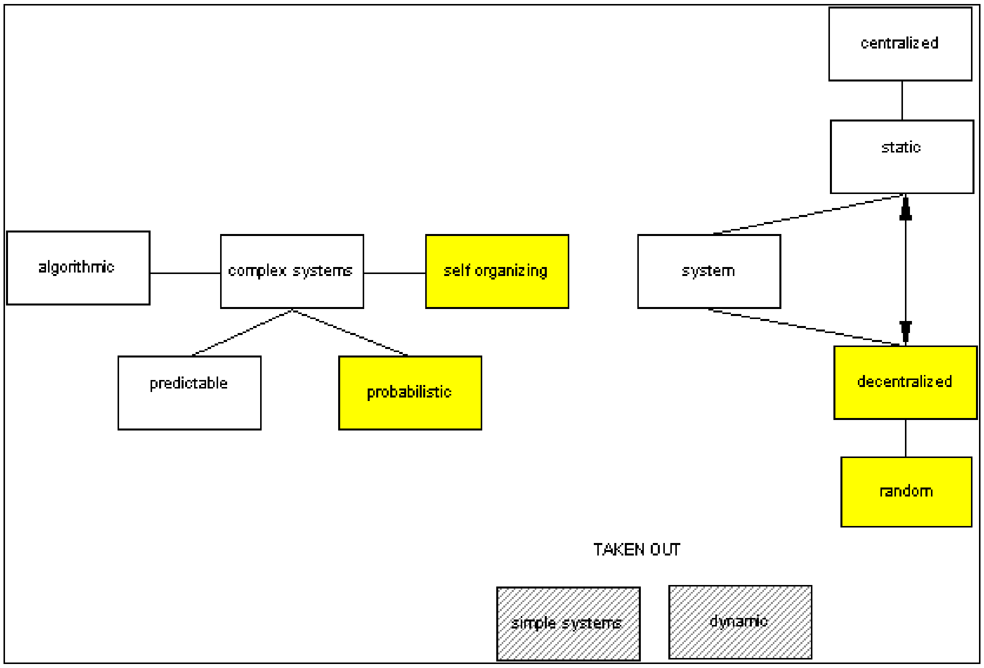
Walter: (talking over I) I remember, I remember like the... The clusters there, of like the ants, like the ant hills or whatever... 255

Walter: Like that, that I could... I felt that that was an emergent property, because you know that they're going to, form clusters. But you don't know exactly where they're going to be, exactly, because it's still governed by probability or whatnot, but you know that they're going to occur. 259

Norman (Figure 12) like Walter also had a somewhat hierarchical structure to his concept map. Like Sidney, he too displayed the radiating type of relationship of concepts suggesting a less sophisticated relationship between terms.

Two important differences in Norman’s map are (1) is removal of the term “random”, and (2) his failure to add the term “emergent”. Unlike other students within this classification, Norman did not integrate the purpose of the study with this activity. It appears that from a metaconceptual point of view he experienced the simulations as completely separate from the assessment activities. This point will become clearer when looking at the results of the outcome measures.

Although Norman experienced a substantial change between session 2 and 6 as evidenced in the changes in his maps; and although he demonstrated a moderate understanding of emergent causal processes through the connection of the terms “self-organization” and “probabilistic” behavior; he also appeared to be seriously hampered by his component beliefs. Looking at his concept map the placement of the term centralized directly beneath self-organization without additional qualification, and placing decentralized under algorithmic are tell tail signs of the conceptual struggle describe earlier. This is additional evidence of the strong clockwork component beliefs guiding his thoughts and limiting his understanding of emergent causal processes.



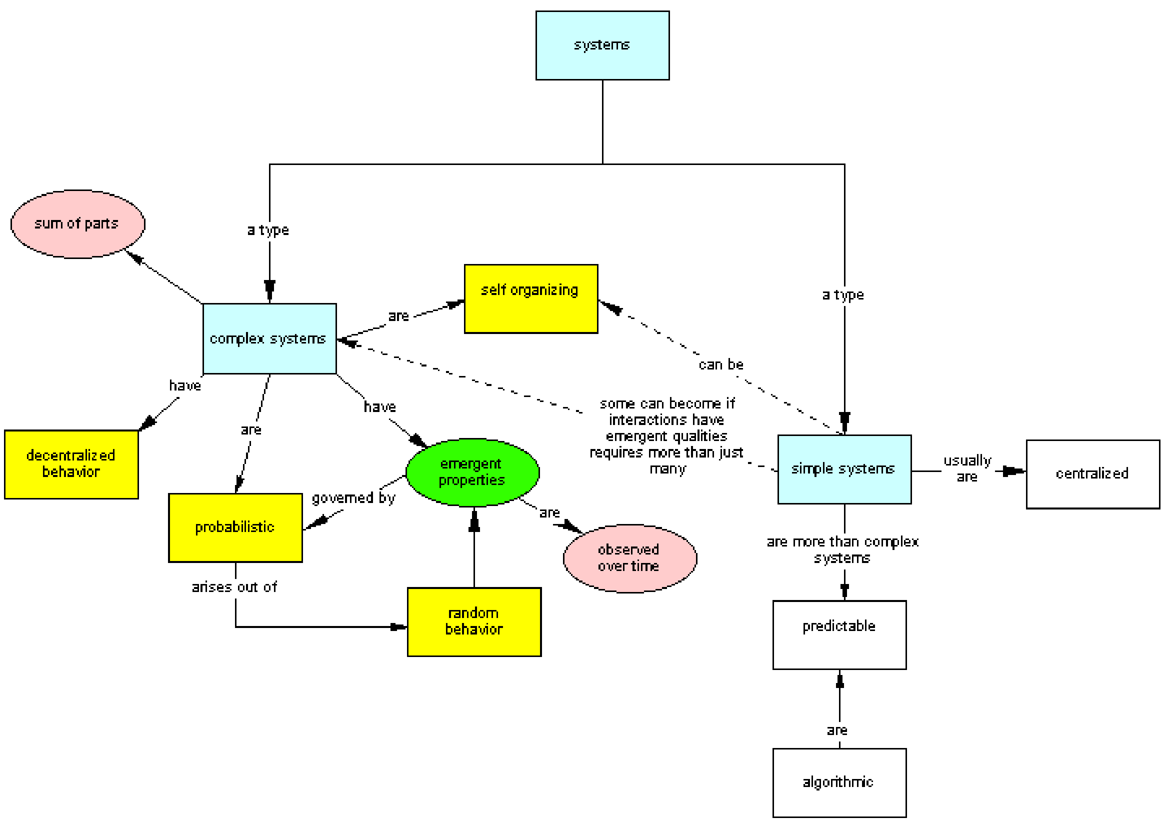


Figure 10. Walter's concept map - Session 2 & 6 (session 2 upper, session 6 lower).

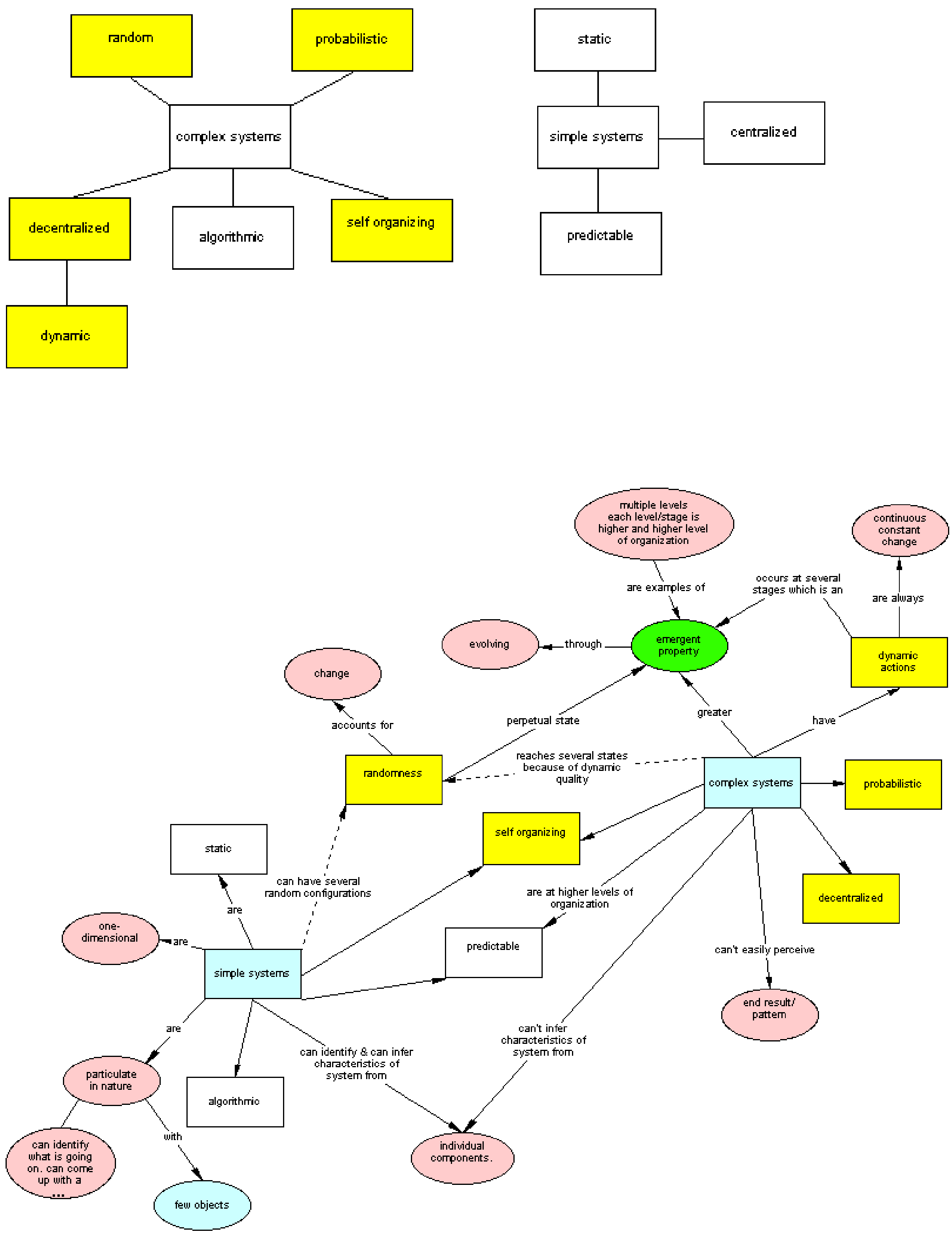
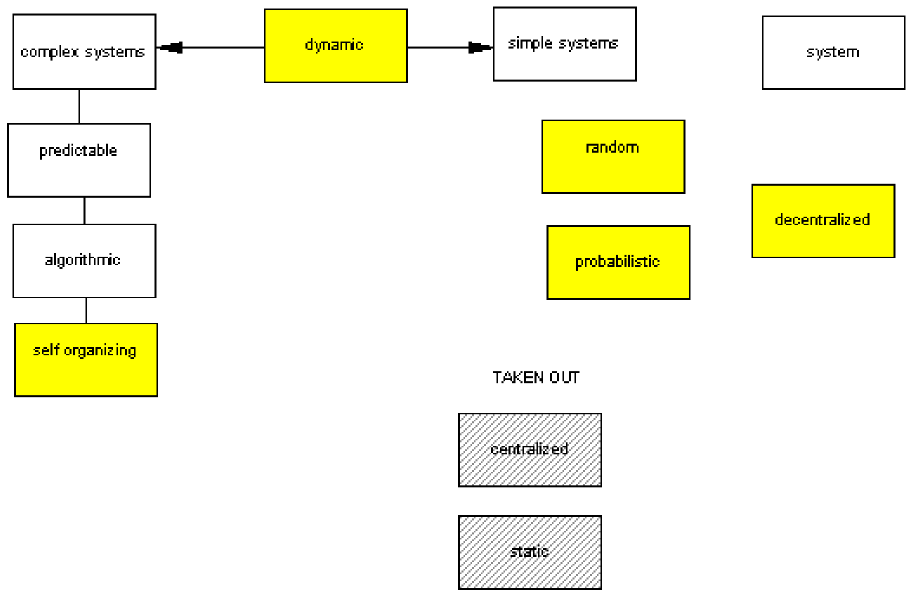


Figure 11. Sidney's concept map - Session 2 & 6 (session 2 upper, session 6 lower).



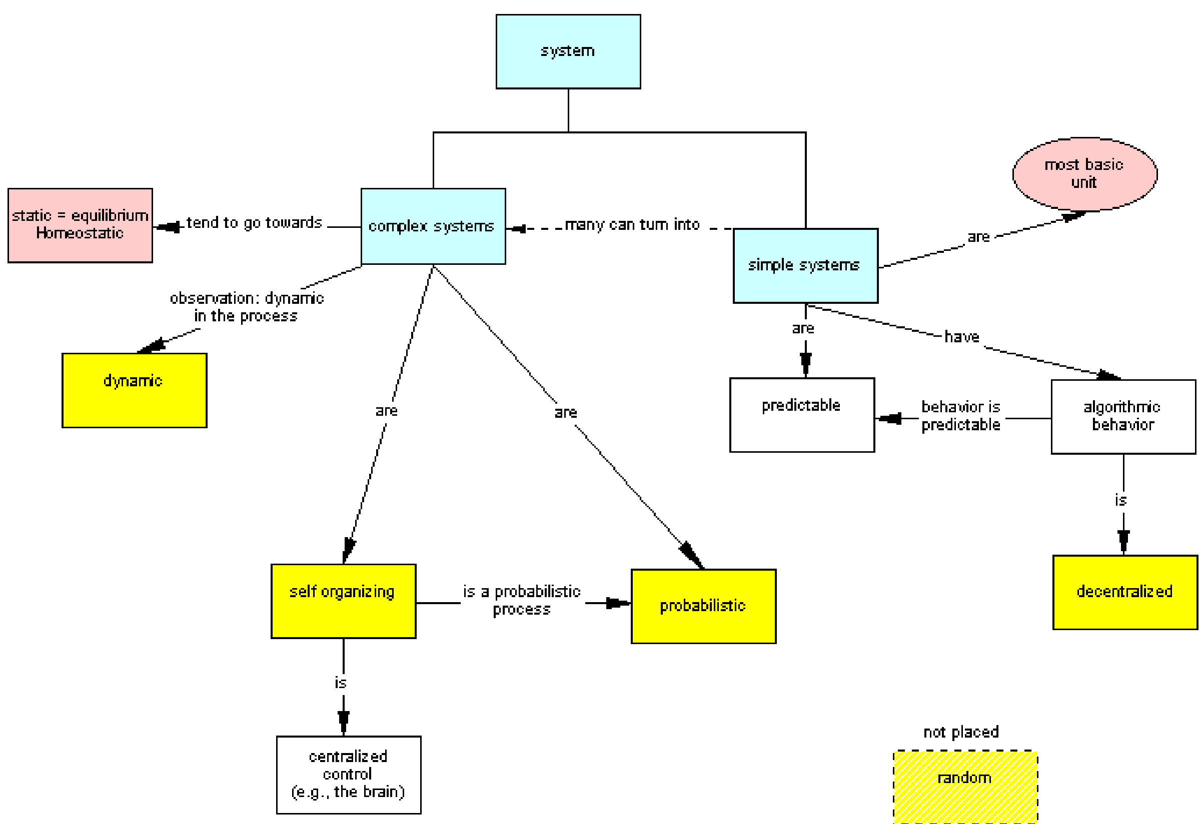


Figure 12. Norman's concept map - Session 2 & 6 (session 2 upper, session 6 lower).

4.3.3 Group 3, “Novice” Understanding of Complex Systems

Penny and Emilie fall into the third classification of maps. Both students constructed very different types of representations, however both are at a “novice” level of understanding. Whereas Emilie (see Figure 14) drew a hierarchical map, Penny (see Figure 13) created a chain-like map quasi-procedural type map (Ruiz-Primo and Shavelson, 1996). Therefore it may be that Penny viewed this knowledge in more of a procedural fashion or maybe she felt safer sticking with a simple explanation because she did not know how to describe her developing understanding. This interpretation is consistent with DeSimoni and Schmid’s (in press) findings that students’ fall into one of three classification of mapping strategies. From their description it appears that Penny could be labeled a “safe player” whereas Emilie would be best described as “limited processor”. Both these types of strategies may have affected how much either student acquired knowledge from this concept mapping process. For instances, even with the direct intervening of the coach, and although she was dissatisfied with her map, Penny did not change the arrangement of her map. This “safe player” strategy may be due to her understanding of the concepts. As demonstrated elsewhere, Penny found the content itself to be challenging.

Penny: Um, it's, here it says it has self-organization...	973
I: Mm-hm.	975
Penny: ...which, is part of both of them?	977
I: It's part of both of them?	979
Penny: Yeah.	981
I: OK, both of...?	983
Penny: Both simple and complex.	985
I: OK.	
Penny: Which is, I think, what I thought before.	989
I: Mm-hm.	991
Penny: First I think I put it under, simple, but then I changed it.	993
(pause 7s)	995
I: Yeah. (pause 18s; papers ruffling in background)	999
I: Yeah this was the last... OK, you have self-organization... Yeah. I: First you had a sort of linear description...	1001
Penny: Yeah.	1003

I: ...with self-organization as simple, and then you went to, you re-drew it, and you had, self-organizing system connected to complex, and simple. And with this thing... 1005

(pause 4s)

1007

Penny: Um... Well this part for sure. But uh, after seeing different, uh, systems, and how they're not linear, I don't know if I should have put them linearly, even though, it makes sense... 1009

I: Mm-hm. 1011

Penny: ...to me, right now, but... Somehow I think, it should be arranged, but I don't know how. 1013

I: How? OK. 1015

Penny: I don't, like... (pause 8s) 1019

I: It's OK, you know, it's just it was... 1021

Penny: No, it, it still makes sense. 1023

I: OK... 1025

Penny: Well because like, when I put it this way, it doesn't mean that, this whole stack that follows that, it just means that they're all, related. 1027

<u>Emilie's map</u>. Examining Emilie's concept map, we also see a great deal of difficulty in the development of her understanding. Her transcript data indicated that her efforts to create a hierarchical structure reflected her clockwork mental model. It appeared that the general framework used to understand the simulations, what may be better described as an "analogy", appeared to be a non- emergent hierarchy (i.e., a political system). If fact she explicitly states this on several separate occasions.

Session 4

Emilie: Now I'm thinking more than one, like, one sort of system. Like, sort of like mini-system I was going to say, in the um, in the, let's say, in the bigger system. Then you have like different kind of structures in each system, kind of like... Hm. Kind of like politics.

Later in session 4 she again states:

Emilie: I don't know, um, like you got, what do you got you have Liberals, you've got Conservatives, and they're all in the same kind of, you know they're all in Ottawa, and, you know, they're always obviously fighting all the time, but I think it's they're all part of the same system which is politics, and you know, the way Canada is going to work, or...

Again during session 5 she says:

Emilie: I don't know, I guess I always understood the um, I don't know, the social world better, like um... I guess, it wouldn't only be political but... So now, you know, I don't know when I talk about systems and complex systems and then simple systems driving from that, of a system I think more like a country, a complex system, would be kind of smaller than this huge system, it would be more like uh, provinces as I said.
And then each system, sorry, each simple system, would be um, would be like a town. Or you know either, oh now, that I think rural and urban, I should maybe still do that. Um... 481

Finally, there appeared to be a robust reductive clockwork component belief in Emilie's thinking. In fact, it almost seemed to be a conscience resistance to "seeing" the evidence as demonstrated by the simulations as well as the coach's prompts. Her reaction is best described by Chinn and Brewer's (1993) description of the fifth reaction to anomalous data "reinterpret but retain original theory". Therefore in addition to trying to understand the content knowledge she was also challenged by her component belief. Below is an excerpt from her discussion in which she insists that systems can be reduced. Of all nine participants, Emilie was the only one who appeared to demonstrate this level of conceptual conflict and this level of entrenched clockwork belief.

Emilie: It's just... it's not more confusing, it's just what I.. I don't know. It's just that, well I still stick to what I say earlier, is that I would not think of let's say one making, of like decomposing a complex system, I would not think of it, like there's only one way to do that. Like in politics I see this totally in a different way, or... Not totally, but I don't know, still kind of a different way, I would put... 954

I: OK, you've mentioned that a couple of times. Decomposing a complex system. Do you think you can really do that effectively, of should it be the other way around? 956

Emilie: What do you mean, the other way around? Like... Like the way you analyze a new sentence, is that what... 958

I: Yeah, the way that you would try to understand it. 960

Emilie: I think you would more like go and put things back together, and sort of, you know, separate them, because you end up with... 962

(pause 10sec)

I: Think about your body? If we broke it apart would you be the same? 972

Emilie: Sure, why not? (laughs) I don't know... 974

I: Could we put you back together again, afterward? 976

Emilie: Probably not. 978

I: Yeah, and why is that? 980

(pause 5s) 982

Emilie: I don't know, probably because you're used to some certain structure, and the way they are, and like, the idea of putting them back together, we'll not put them back in the right place, or they'll not necessarily know how to function. And, I don't know, the sort of so-called human body. I don't know.
Because I'm, I'm thinking of like all those things they do with, I don't know, like take the example of fish, like fish that are, or I don't know some animals, they take them into the zoo, in order for them to reproduce, and, like, they've never really tried to put them back in their environment, like, if they, if they've been born in the zoo, and they've been fed by a human, and if they've never really caught their own prey, and then if you want to put them back into nature, they would not necessarily know how to survive or how to react. So... I don't know. I'm thinking that this would probably not... 986

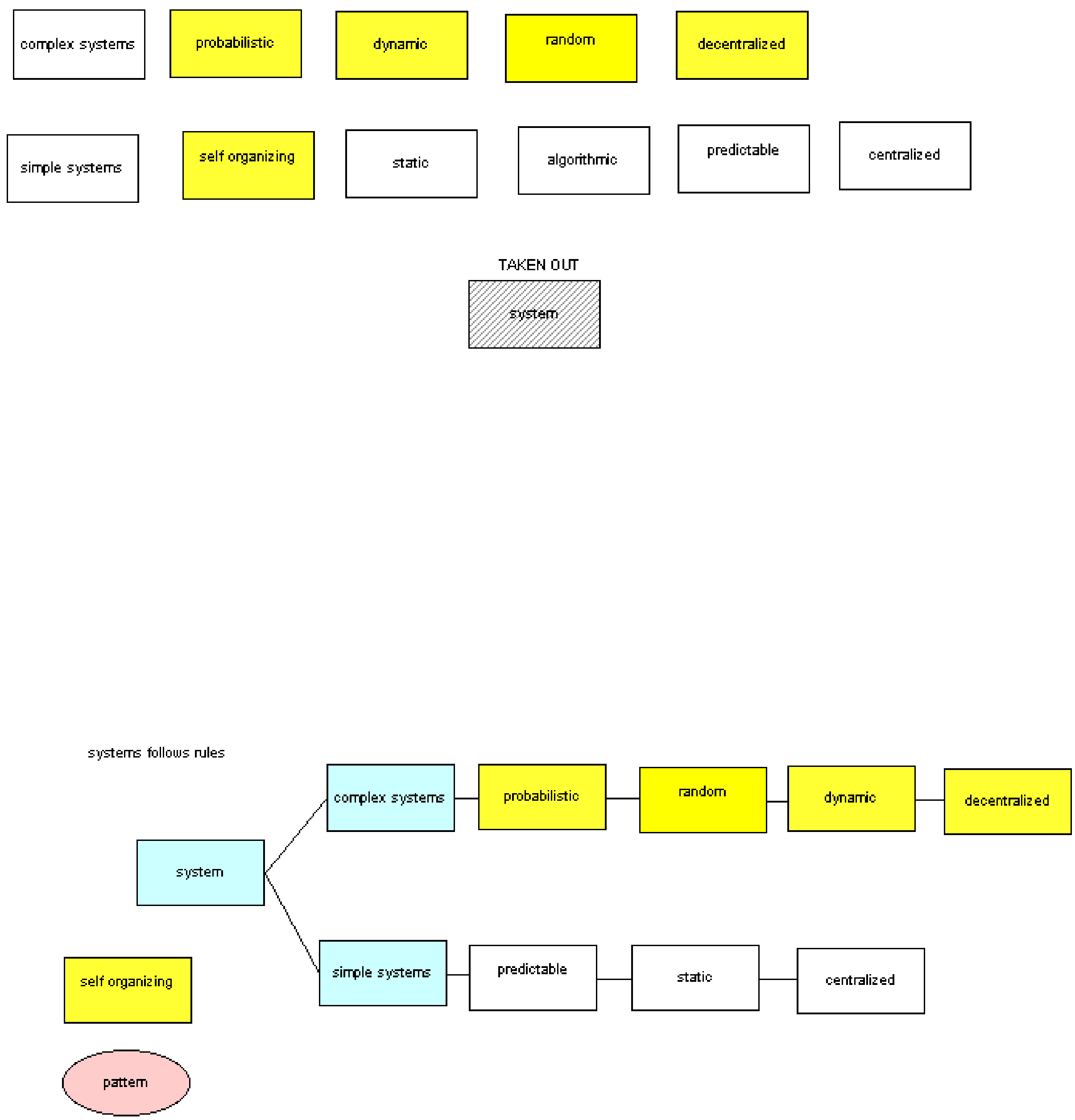


Figure 13. Penny's concept map - Session 2 & 6 (session 2 upper, session 6 lower).

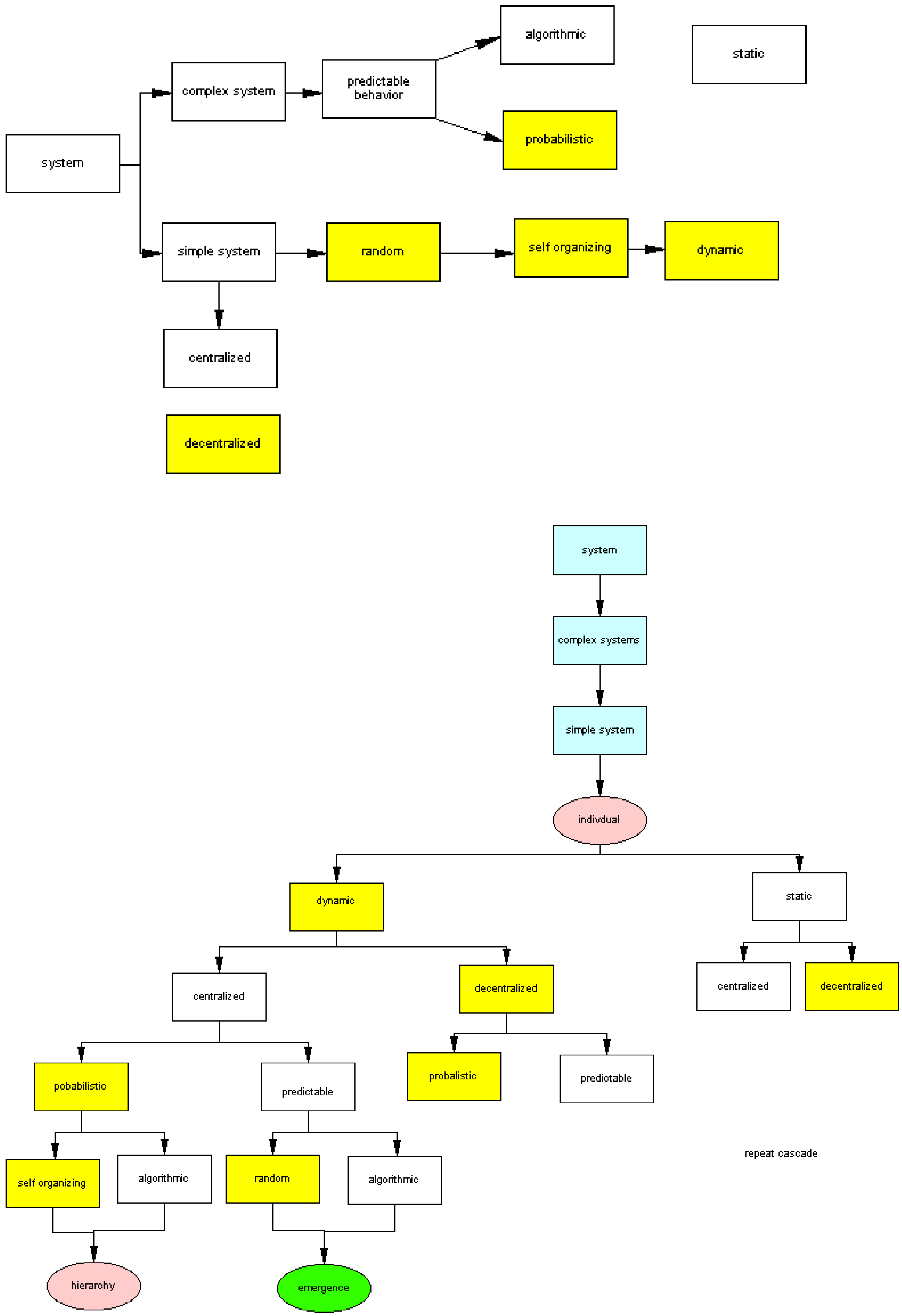


Figure 14. Emilie's concept map - Session 2 & 6 (session 2 upper, session 6 lower).

4.3.4 Summary on Results of the Concept Mapping Activity

The concept maps show that most students experienced considerable development in their understanding of the concepts regarding complex systems behaviors. Over half were able to construct relationships between the key concept nodes of “complex system”, “self-organization” and “probabilistic behaviors”. This understanding was also seen in the transcript data (evidence discussed in the next section). Four students were able to identify the importance of the concept “emergent behaviors” by deciding on their own to incorporate the term into their concept maps. This marked a significant development in their understanding of how complex systems are related to emergent causal processes. This was noteworthy because some of these changes to their maps were made at the final interview, one year after the end of the intervention (i.e., session 6).

Lastly, one other important observation was made from these concept maps. It appeared that the concepts of “random” and “predictability” were particularly problematic for some students – a common finding in the literature (e.g., Wilensky 2001). In fact, Norman completely removed the term “random” from his final map; and links “predictable” to simple systems. For Sidney, it appeared that although he had developed a good understanding of randomness as accounting for change he still did not link it to the probabilistic nature of complex systems. This was surprising given his transcript data indicated otherwise. However, he demonstrated a more sophisticated understanding of “predictability” by annotating the link with the conditional statement of “at a higher level”. This suggested that his understanding of “predictability” was more in line with the greater stability exhibited by higher levels meta-agents as demonstrated by the simulations. On the other hand, Greg, Sam and Walter, from their maps appeared to understand the important relationships between “randomness” and “predictability”. However, this should not suggest that it was without a certain degree of cognitive struggle as demonstrated on several occasions during the intervention by their metacognitive discourse (for examples see excerpts in Appendix D).

DISCUSSION

5.1 Limitations to Understanding of Ontological Category of Emergence

Chi (2000) proposed that there are three major limitations (inter-related barriers) to the understanding of the ontological category of emergence: (1) assignment of micro level behavior as linear, (2) lack of consideration of local interactions between agents, and (3) a lack of understanding that macro level emergence is the result of collective interactions of agents and environment – "interaction in dynamic collection". Additionally, the literature concerning science misconceptions identified six ontological barriers[12] that in many ways overlap with those described above. Chi (Slotta & Chi, 1999) further proposed that ontological training could remove these barriers.

The results of this current study shows that not all of these identified barriers are equally challenging, in fact the instructional intervention was able to change at least two of the six subcategory dimensions. However, the evidence also suggests that two of these ontological barriers are not affected by the chosen intervention, and at least one of these appears to form a firmly entrenched belief that requires special conditions before it may be addressed.

An analysis of the change in mental models (Table 6) suggested that the ontological training produced four patterns of conceptual change. Four students (Monique, Emilie, Penny, and Walter) exhibited a novice emergent mental model, one student (Norman) exhibited a synthetic 2A mental model, three students (Mitch, Sidney and Sam) exhibited a synthetic 1A mental model, and one student (Greg) exhibited an emergent mental model. This pattern is very similar to the pattern of acquisition of conceptual understanding of mental models (Figure 4). The only difference is that Walter had a higher level of content knowledge (Complex Systems) that was revealed by his explanatory framework or mental model.

---

[12] 1) Isomorphic behavior of both micro and micro levels behaviors (reductive ontology) to emergent aggregation behaviors (non-reductive ontology);
2) Centralized control to distributed or decentralized control (decentralized control);
3) Linear causal explanation of macro-level behavior from micro-level interactions (i.e., additive, linear) to multiple nonlinear causal explanations (nonlinear effects).
4) Determinacy to indeterminacy (random actions);
5) Intentionality (i.e., teleological) to stochastic causes (i.e., probabilistic causes);
6) Static processes (i.e., beginning-end processes) to dynamic homeostatic behaviors (dynamic self-organizing nature).

5.1.1 Multiple Levels of Organization

Understanding the concept of "multiple-levels" requires that the learner is able to appreciate the different behaviors exhibited by the "agent" (as an independent unit within the system), as well as part of the "meta-agent" (the emergent aggregate unit at a higher level of organization within the system). The literature (Duit et al., 2001; Jacobson & Archodidou, 2000; Penner, 2000) and Chi (2000) suggest that grasping this concept is not an easy task.

All students appeared to acquire and use the concept "multiple levels of organization". However, their reference to it decreased over the three simulations. Whereas 49% of the total observations made during the Slime simulation referred to multiple levels of organization, during the FreeGas and Wolf-Sheep simulations the percentage of statements coded to this concept falls to 35% and 32.5% respectively. The three simulations offered great affordances for this concept.

- The Slime simulation, as an example of a tightly coupled organized complexity model, exhibits emergence through a physical and perceptual coming together (aggregation) of agents. Once these agents are in their aggregate form, they display perceptually different behaviors.
- The FreeGas simulation, as an example of a random disorganized, dissipative, complexity model, exhibits emergent aggregation at more abstract levels. For instance, the meta-agent behaviors may be seen at the statistical level of probabilities where larger populations of molecules produce more stable and predictable results. Or the meta-agent may be understood at the mathematical symbolic level in which the equation $Pv=nRT$ operates to relate different pressure, or temperature values (interpreted as energy and speed on the FreeGas simulation) depending on the number of molecules and their initial velocities.
- The Wolf-Sheep simulation is an example of a complex system somewhere in-between these two other types of complexity models. Like the FreeGas simulation, it too requires the learner to be cognizant of a somewhat abstract level of organization (i.e., the oscillating sine wave population variation which is the graphical interpretation of the symbol level Lotka-Volterra equation: $dn_1/dt = n_1(b-k_1n_2$ and $dn_2/dt = n_2(k_2n_1-d)$ used in predator prey interactions).

Differences of understanding of the concept of multiple-levels may be due to at least three different factors: conceptual, perceptual and individual differences.

Conceptual. The results in this study suggest that when learning about multiple-levels of organization (emergent levels) students were less likely to understand this concept when working with dissipative models of complexity than when they were working with tightly coupled organized models of complexity. One explanation might be that dissipative systems are a more difficult subcategory of the emergent causal processes. Therefore they may require more time, cognitive effort, scaffolding, or a certain type of "ontological readiness" to facilitate understanding. In fact, the observed difficulty of understanding the concept of "multiple levels of organization" from dissipative models correlates with the literature regarding misconceptions and difficulties in understanding dissipative systems: diffusion, gas laws and equilibrium in chemistry (Wilson, 1998); electricity in physics (Chi, Feltovich, & Glaser, 1981; White, 1993); and diffusion, and osmosis in the biological sciences (Odom, 1995; Settlage, 1994).

By contrast, students appeared to have little difficulty with understanding "multiple levels" from the Slime simulation. It may be that its representation of tightly coupled organizing complexity is more accessible to novice learners. After all we experience such on a daily basis (e.g., social groupings like families and friends, neighbourhoods, schools, etc.). Furthermore they are more easily identified in nature, colonies of ants, flocks of birds, sheep, for example.

Perceptual. Whereas the Slime simulation produces a visible clustering of agents (mould) into the higher-level meta-agent (colony), there are no such visual observations in the other two simulations. There the learner is dependent on other visual devices such as window displays of graphs. Learning such abstract meta-level states may impose a greater cognitive workload or demand a higher level of knowledge in order to understand the different representation. From the multimedia literature we are told that visual representations such as text and images presented simultaneously are more taxing to working memory (Mayer & Moreno, 1998; Clarke & Paivio, 1991). It is therefore possible that a similar cognitive overload is at work when students attempt to decode both the animation of the agents' lower level behaviors and the graphs of the systems' higher level behavior.

Individual. The evidence shows that excerpts from Greg, Mitch, and Sidney, the students required appreciable amounts of scaffolding to interpret the output graphs being produced during the FreeGas simulation. This suggests that this type of simulation requires more time-on-task, and more cognitive scaffolding in order for learners with weaker science backgrounds to gain the full benefit of learning about the emergent processes represented as well as the behaviors of dissipative systems at the level of the individual.

5.1.2 Local Interactions

The results presented in this report suggest that the concept "local interactions" is the most susceptible to acquisition from the three simulations. In essence, all students appeared to show substantial gains in awareness as they proceeded through the instructional sessions, and all students appeared to transfer some of this knowledge to their explanation of the ontologically analogous problems. Slime, FreeGas, and Wolf-Sheep simulations produced a steady increase in awareness to (and possibly learning of) the concept of "local interactions" (22%, 25%, and 32.5%, respectively). This small but steady increase is the type of result one would expect from an instructional intervention, which produced a significant improvement in learning. However, the changes observed in this study also suggest several levels of interactions between the intervention, the concepts and the students.

Understanding of "local interactions" arising from the simulations may be explained as a consequence of two features. Firstly, the surface level visual cueing of tightly coupled interactions and aggregation that were seen in the Slime simulation. Secondly, the causal chain cueing in the FreeGas, and Wolf-Sheep simulations. In short, the collision of one molecule in the dissipative loosely coupled system still created a change in trajectory which would result in another collision and so on. Meanwhile, the ecological systems' food chain of "sheep eat the grass... wolves eat the sheep", with moderately coupled interactions, was also easily apparent.

In addition, the increase in observations of local interactions from one session to the next could also be explained as a result of the students' improving ability to observe emergent processes at a more structural level (i.e., readout strategies). For instances, students "saw" more than the direct interactions between the wolf, sheep and grass, and described the indirect relationships in this food chain. Although not all students were able to identify the nonlinear effects of these local interactions, they all spent considerable time trying to explain how the grass affected the overall balance and survival of the systems.

Other studies looking at learning of complex systems concepts have also reported similar results. Using conceptually similar, although very different, measures and media ("talus slope" and “Life” simulation), Penner (2000) reported that students increasingly recognized that micro level interactions were important to the systems’ behavior.

The results suggest that awareness of the concept of “local interactions” was not only related to the opportunities provided by the simulations, but also by the abilities of individual students. In fact, Greg and Mitch both demonstrated high levels of awareness of this concept during their engagement with the FreeGas simulation (61%, and 43% respectively). Both Greg and Mitch belonged to a cohort of high academic achievers and both had above average scores in their science courses, particularly in college introductory level Physics. Thus, there may be an interaction between the student’s domain knowledge and their ability to perceive these collisions as interactions of objects that although inanimate, engage in a flow of energy through the system; therefore, displaying behaviors that could be described using a general explanatory model, which could also apply to systems as diverse as slime mould. In fact, the successful results reported by the StarLogo researchers (e.g., Resnick, 1994, Wilensky, 1999, 2001) may well be due to this ability to “think like a turtle”. In other words, most of the research up to now has been conducted with younger children. It may well be that there is a fine line between anthropomorphizing and the ability to think at the level of the individual agent. It may be that Greg and Mitch, not children but advance-level science students, could think at this level and could “see” the collision of gas molecules as interactions and all that they entail (i.e., flows of energy through the system, etc.). Further investigation into this possible relationship is required.

Additionally, looking at the specific fine grain differences, Mitch as early as session one, began describing the interactions of agents (the slime mould in this instances) using predicates that better describe molecules (e.g., “collisions”, “collide”, “hit across”, to list a few). This may explain his underlying understanding of the importance of local interactions and consequently this dimension of the emergent ontological category.

5.1.3 Nonlinearity and Random Behavior

Nonlinearity and randomness are examples of two Complex Systems concepts that presented difficulty to many students. Although the experts identified at least three different types of nonlinearity in the three simulations (saturation, positive feedback loop, and negative feedback loop), students exhibited little awareness of this concept in the three simulations. In fact no student demonstrated a grasp of this concept. However as some students did refer to it in the FreeGas and Wolf-Sheep simulations, this knowledge was available. In addition to the weak opportunities provided by the simulations for learning this concept, there are possibly ontologically-based restrictions. In fact, Chi (2000) identifies this as the first of the three ontological barriers based on novice learners' predisposition to explain micro-level behaviors as linear in nature and leading to linear predictable outcomes (labeled item #1).

Therefore, is it a question of StarLogo's affordances for teaching this concept or is it the concept itself? As stated before, Penner's (2000) study examining the use of a "talus slope" tool and the "Life" simulation, reports that students provided evidence of some recognition that small micro-level changes can have significant macro-level effects. Unfortunately he does not provide details on the number of students and percentage of change therefore it is difficult to compare these tools to StarLogo models. However, his study suggests that other tools may be more successful representations of this phenomenon. Therefore the removal of this barrier through training is still an unanswered question. What can be stated from the evidence is that students like Greg, Mitch, and Sidney to a lesser degree, who are conceptually prepared and understand other aspects of emergent behaviors are able to appreciate the impact of nonlinearity in the systems created in StarLogo. This effect is amplified through the coach who was able to prompt for more elaboration and metacognitive explanations.

One of the ontological barriers **not** identified by Chi (2000) is the attribution of causal determinacy (i.e., difficulty in acquiring the concept of random actions). This current study shows that, possibly because of weak affordances of the models for learning this concept, students experienced difficulty with the notion of randomness. Klopfer and Um (2002) in a study of fifth and seventh grade students using StarLogo in a scaffolded learning environment called "Adventures in Modeling" also demonstrated that students experienced difficulties with learning the concept of random events; although in the latter portion of their 14 sessions intervention, students were able to grasp this concept.

The evidence from the study reported here is that all students at some level were challenged by randomness. In fact, it was the main stumbling block for Greg who otherwise acquired an understanding of all the emergent causal processes without exceptional cognitive struggle. What this suggests perhaps is that even though students accept the randomness of some happenings, as indicated in their answers to the question about ants foraging, at a deeper level they struggle to accept the lack of some means of predicting future outcomes (even by infinitesimally small or remote means). This deep level understanding is further confounded by the limitations of the programmed environment of the simulations, which indeed may confirm beliefs that there is some level of predictability because random number generations machines are behind these calculations. In fact, this is the level of discussion that Greg, Mitch and Sidney all at some point conducted with the coach.

How then did any of the students show signs of acquiring a deeper level understanding of this concept? The evidence suggests that Greg was the only student to describe random actions at the deeper level of understanding as an element of true in causal determinacy and “noise”. He appeared to accomplish this as a consequence of both cognitive scaffolding and his domain knowledge. In essence, during the final interview session, one year after the intervention, Greg was asked to explain his concept map. During this discussion, he elaborated on the role played by random actions in the behavior of systems. This required him to reflect and in doing so he referenced his course work from biology and how the “noise” of random events creates the “possibilities” of the future states.

The attribution of causal determinacy is a key obstacle to understanding emergent causal process for most learners. This arises either because of the learners’ component beliefs, as in the instantiation of the case study Norman, or because of the confounding of concept and programming limitations as demonstrated by Sidney, Mitch and overcome by Greg. This may come as no surprise to those investigating the cognitive processes involved in reasoning about uncertainty (e.g., Shauhnessy, 1992; Tversky & Kahneman, 1974). Metz (1998) points to the spurious causal attributions that result from misunderstanding of randomness and probability . What is surprising is that this same barrier also may account for a major difficulty in learning emergent causal processes such as evolution. This contention is supported by research from Zaïm-Idrissi, Désautels, and Larochelle (1993) in which, working with 15 graduate level biology students (master’s level), they concluded that the majority of the students held deterministic forms of reasoning about the topic of evolution. These authors uncovered several inconsistencies in the belief systems of the study’s participants, primarily, the conflict between deterministic and probabilistic reasoning.

Therefore, it is possible that this causal determinacy attribution may be one of the most widely interconnected beliefs that affect other related beliefs such as probabilistic causes, and even decentralized control. It may well fit Chinn and Brewer’s (1993) description of the evidentiary supporting schema. They state: “It appears, then, that well-developed schemas are not necessarily entrenched. The key is whether the schema is also embedded in evidentiary support and is used to support a wide range of other theories and observations that the person believes" (p. 17). Future research is required to try and untangle the possible confounding of the simulations' weak affordances and the students' ontological belief about randomness.

5.2 Triangulation of Data Sources

Triangulation of data collected from different sources is a recommended practice used primarily, but not exclusively, in qualitative research studies to establish validity (i.e., trustworthiness and authenticity). This procedure of drawing together and comparing data collected from different techniques addresses the threat of experimental bias, which may be inherent in particular data sources, investigator bias, and methods (Creswell, 1994). The benefit of employing a mixed methods research design as in this current study was that it provided these requisite differences between data sources.

In order to conduct the triangulation, we brought together the results from the four main data collection instruments used in the case study: (1) Emergent Framework Mental Model (EFMM), (2) Clockwork Mental Model (CWMM), (3) Complex Systems Taxonomy (CST), and (4) concept maps. We also included the students' scores on the Nelson Denny Comprehension and Reading test, as well as their GPA.

Table 12 shows the Pearson Product correlations between the above data sets. They indicate that, as expected, there is a significant strong positive correlation between the students' understanding of complex systems as indicated by their scores on the EFMM, CST, and concept maps and their academic ability as indicated by their GPA and two scores of the Nelson Denny. On the other hand, there was no significant relationship or a significant negative relationship between students' scores on the CWMM and all other measures.

Table 12. Pearson product correlations between different data sources.

	CST	EFMM	CWMM	GPA	ND (VOC)	ND (COMP)	CONCEPT
EFMM	.946*	1.000	-.805*	.897*	.949	.908*	.826*
CWMM	-.846*	-.805*	1.000	-.779*	-.717	-.725	-.632
GPA	.955*	.897*	-.779*	1.000	.788*	.693	.874*
ND (VOC)	.895*	.949*	-.717	.788*	1.000	.974*	.815*
ND (COMP)	.843*	.908*	-.725	.693	.974*	1.000	.794*
CONCEPT	.910*	.826*	-.632	.874*	.815*	.794*	1.000

$df_r = 7$, * indicate significance at least at $\alpha = 0.05$,

Finally, the three measures of students' understanding of complex systems (CST, EFMM, and concept maps) are all significantly correlated (at least at $\alpha = .01$). Moreover, 84% of the variability in students' scores on the concept mapping score are predicted by their scores on the CST and EMFF ($R^2 = .84$). These consistent correlation results supports the claim that the data collection instruments were measuring the same phenomena. This statistical triangulation of the data sources adds trustworthiness to the data analysis methods.

In conclusion, a non-statistical comparison of the data sources describes a similar overlapping of results. Specifically, the classifications of student experiences identified from the CST results (*ECP Identifiers*), as well as the concept map results (*Understanding of ECP Relationships*), lastly the OMMT results (*EFMMs Producers*). These three data sets show the same students, more or less, classified as equivalent levels of understanding across these measures (see Table 13). This consistent pattern is another way of demonstrating the triangulation the three data sets collected in this study.

Table 13. Classification of case study students* across the three data sources.

Classification	Descriptions from Data Sources		
	ECP Identifier	*ECP Relationships*	*EFMMs Producers*
Sophisticated	Greg	Greg	Greg Mitch Sam
High moderate	Mitch Sam Sidney	Mitch Sam	
Moderate	Walter Norman	Walter Sidney Norman	Sidney
Novice	Penny Emilie	Penny Emilie	Walter Norman Penny Emilie

(N.B. EFMM Producers based on final posttest results).
* Only students for whom we had full data were included.

5.3 Educational Implications

Three main educational implications can be drawn from this study. The first implication concerns the ease with which most students can acquire components of emergent framework mental models from short-term interventions. These students showed considerable gains in learning about emergent causal processes and were able to apply some of these concepts to transfer problems.

The second implication refers to the need for a greater understanding of emergent causal processes by curriculum developers (e.g., instructional designers) and teachers so that they are more aware of the many opportunities to apply this knowledge. In fact, many teachers and curriculum developers lack an appreciation of the constraints imposed by their own linear, reductive thinking. Therefore, part of the challenge will be to convey to the educational institutions that prepare teachers, instructional and educational technologists an understanding of the benefits of emergent causal thinking as a general problem-solving application framework. Additionally, until recently there has been a lack of representational tools to readily convey emergent processes as demonstrated by complex systems and thereby provide the necessary scaffolding for learning these concepts. While these tools are making their way into the educational system, there is a need to develop the easily accessible curricula topics that demonstrate complex systems behaviors (e.g., respiration, and cardiovascular circulation in the health sciences, the behavior of geological and ecological systems in the natural sciences).

The third implication is that this alternative explanatory framework may be beneficial for all disciplines not just science. If students are better able to explain the social, political, and economic interactions they encounter with more than a linear perspective they may in fact do a better job of understanding the unpredictable, and probabilistic nature of many of these phenomena. In fact, some proponents of complexity theory suggest that interpreting political interactions in terms of complex adaptive systems would have foretold of recent world events better than the current linear, reductive models.

5.4 Future Directions

This study shows us that it is possible to use newly available representational technologies to create specific content models to teach general knowledge of emergent causal processes. However, affordances of different models for learning specific emergent causal processes are interwoven with the content domain knowledge and representational characteristics. If we are to make better use of existing models, we need to know more about their capabilities. Therefore, these potentially confounding interactions of content and model representation of emergent process need to be closely monitored. Furthermore, there is a need to better comprehend the affordances of various media, not just computer simulations, for helping students construct the mental representations necessary to understand important complex systems knowledge of relevance to the natural and social sciences.

CONCLUSION

In conclusion, the results of this study provide strong support that the ontological training facilitated the creation of emergent framework mental models (EFMMs). This conclusion is supported by the evidence that most students acquire at least four of the Complex-Systems concepts

The affordances for learning aspects of emergent causal processes offered by the multi-agent models/simulations are highly related to the type of complex system represented and also to the students' background understanding of science. In particular more students had difficulty learning with representations (simulations) of dissipative system complexity compared to those using representations of tightly coupled organization models of complexity.

Conceptual change requires not only robust conceptual representations (e.g., models that can be used as analogies) but also metacognitive scaffolding and ongoing metaconceptual prompts during the instructional phase. Once initiated (i.e., once synthetic mental models are created), maturation over time and experience with complementary domain curricula appear to have positive effects on the development of more elaborated emergent framework mental models.

REFERENCES

Acton, W.H., Johnson, P.J., & Goldsmith, T.E. (1994). Structural knowledge assessment: Comparison of referent structures. *Journal of Educational Psychology*, 86, 303-311.

Ahn, W.(1998). Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality. *Cognition*, 69, 135-178.

Anderson, C. & Bishop, B. (1986). *Evolution by natural selection: A teaching module*. (ERIC Document Reproduction Service No.ED269254)

Auyang, S. (1998). *Foundations of Complex-system Theories in Economics, Evolutionary Biology, and Statistical Physics*. Cambridge, UK: Cambridge University Press.

Azevedo, R., Seibert, D. Guthrie, J.T., Cromley, J.G., Wang, H., & Tron, M. (2002). *How do students regulate their learning of complex systems with hypermedia?* Paper presented at the Annual Conference of the American Educational Research Association, New Orleans, LA.

Banerjee, A. C., (1995). Teaching chemical equilibrium and thermodynamics in undergraduate general chemistry classes. *Journal of Chemical Education*,72(10), 879-881.

Barsalou, L. (1989). In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp.). Cambridge, MA: Cambridge University Press.

Bar-Yam, Y. (1997). *Dynamics of complex systems*. Reading, MA: Addison-Wesley.

Bishop, B.A., & Anderson, C.W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching*, 27(5), 415-427.

Bloom, J.W. (2001). Discourse, cognition, and chaotic systems: An examination of students' argument about density. *Journal of the Learning Sciences*, 10(4), 447-492.

Boyd, G.M. (1997). Excavating the emergent levels of cybersystemics. R. Glamtle & G. DeZeeuw, (Eds.), *Problems of Excavating Cybernetics and Systems, Systemica*, 11(1), 29-42.

Brumby, M. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education*, 68(4), 493-503.

Capra, F. (1996). *The Web of Life: A new scientific understanding of living systems*. NY: Anchor Books, Doubleday.

Champagne, A., Gunstone, R., & Klopfer, L. (1985). Effecting changes in cognitive structures among physical science students. In L. West & L. Pines (Eds.), *Cognitive Structure and Conceptual Change.* Orlando, FL: Academic Press.

Chan, C., Burtis, J. & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction*, 15(1), 1-40.

Chi, M.T.H. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), *Cognitive models in science: Minnesota studies in the philosophy of science* (pp. 129-186). Minneapolis, MA: University of Minnesota Press.

Chi, M. T. H. (1993). *Barriers to conceptual change in learning science concepts: A theoretical conjecture*. In W. Kintsch (Ed.), Proceedings of the Fifteenth Annual Cognitive Science Society Conference (pp. 312-317). Hillsdale, NJ: Erlbaum.

Chi, M.T.H. (1997). Qualifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, 6(3), 271-315.

Chi, M.T.H. (2000). *Misunderstanding emergent processes as causal*. Paper presented at the Annual Conference of the American Educational Research Association, April 2000.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates

Chi, M.T.H. & Roscoe, R.D. (2002). The processes and challenges of conceptual change. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change. Issues in theory and practice* (pp. 3-27). Netherlands: Kluwer Academic Publishers.

Chi, M.T.H. & Slotta, J.D. (1993). The ontological coherence of intuitive physics. *Cognition and Instruction*, 10 (2 &3), 249-260.

Chi, M. T. H., Slotta, J. D., & deLeeuw, N. (1994). From things to processes: a theory of conceptual change for learning science concepts. Learning and Instruction, 4, 27-43.

Chinn, C. & Brewer, W. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, 63(1), 1-49.

Clark, J.M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3, 149-210.

Coll, R.K. & Treagust, D.F., (2002). Learners' use of analogy and alternative conceptions for chemical bonding: A cross-age study. *Australian Science Teachers' Journal*, 48(1), p.24-32.

Colella, V. (2001). Participatory simulations: Building collaborative understanding through immersive dynamic modeling. *Journal of the Learning Sciences*, 9(4), 471-500.

Creswell, J.W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, CA: SAGE Publications

Creswell, J.W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Merrill Prentice Hall.

Demastes, S.S., Good, R.G., & Peebles, P. (1995). Students' conceptual ecologies and the process of conceptual change in evolution. *Science Education*, 79(6), 637-666.

De Simone, C. & Schmid, R.F. (2003). *Concept mapping from the inside out: Understanding learners' processes, activities and experiences*. Manuscript submitted for publication.

diSessa, A.A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10 (2 &3), 105-225.

diSessa, A.A., Sherin, B., (1998).What changes in conceptual change? International Journal of Science Education, 20 (10), 1155-1191.

Driver, R. (1995). Constructivist approaches to science teaching. In L.P. Steffe, & J. Gales (Eds.), *Constructivism in education* (pp. 385-400). Hillsdale, NJ: Lawrence Erlbaum Associates.

Duit, R., Roth, W.M., Kormorek, M. & Wilbers, J. (2001). Fostering conceptual change by analogies – between Scylla and Charybdis. *Learning and Instruction*, 11, 283-303.

Erickson, F. (1986). Qualitative methods in research on teaching. In M.C. Whittrock (Ed.), *Handbook of research on teaching* (3rd ed.). Old Tappan, NJ: Macmillan.

Feltovich, P., Coulsen, R., Spiro, R.., & Adami, J. (1992). *Conceptual understanding and stability and knowledge shield for fending off conceptual change*. Springfield, IL: Southern Illinois University School of Medicine.

Ferrari & Chi, M.T.H. (1998). The nature of naÃ¯ve explanations of natural selection. International Journal of Science Education, 20(10), 1231-1256.

Flood, R. & Carson, E. (1993). *Dealing with complexity: An introduction to the theory and application of systems science*. New York: Plenum Press.

Gentner, D., & Stevens A.L. (1983). *Mental models*. Hillsdale, NJ: Erlbaum Press.

Glaser, B.G. & Strauss, A.L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Publishing Company.

Glynn, S. & Duit, R. (1995). *Learning science meaningfully: Constructing conceptual models*. In Shawn Glynn & Reinders Duit (Eds.) Learning Science in the Schools: Research Reforming Practice (pp.3-34). Mahwah, NJ: Lawrence Erlbaum Associates.

Goel, A., & Chandrasekaran, B. (1989). Functional representation of designs and redesign problem solving. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence* (pp.1388-1394). Los Altos, CA: Morgan Kaufmann.

Guba, E.G., & Lincoln, Y.S. (1981). *Effective evaluation: Improving the usefulness of evaluations results through responsive and naturalistic approaches*. San Francisco, CA: Jossey-Bass.

Hallden, O. (1988). The Evolution of the Species: Pupil Perspectives and School Perspectives. *International Journal of Science Education*, 10 (5), p.541-52.

Hmelo, C.E., Holton, D.L. & Kolodner, J.L. (2000). Designing to learn about complex systems. *Journal of the Learning Sciences*, 9(3), 247-298.

Holland, J.H. (1995). *Hidden order: How adaptation builds complexity*. Reading, MA: Addisoon-Wesley Publishing Co.

Holland, J. H. (1998). *Emergence: From chaos to order*. Reading, MA: Perseus Books.

Jacobson, M.J. (2000). Butterflies, Traffic Jams, & Cheetahs: Problem solving and complex systems. Paper presented at American Educational Research Association annual meeting in Atlanta.

Jacobson, M. J. (2002, April). *Complex systems in education: Integrative conceptual tools and techniques for understanding the education system itself*. Chair and organizer for symposium present at the 83rd Annual Meeting of the American Educational Research Association, New Orleans, LA. (Participants: Uri Wilensky, Jim Kaput, Jay Lemke. Discussant: Nora Sabelli).

Jacobson, M. J., & Archodidou, A. (2000). The design of hypermedia tools for learning: fostering conceptual change and transfer of complex scientific knowledge. The Journal of the Learning Sciences, 9(2), 149-199.

Jensen, M. & Finley, F. (1996). Changes in students' understanding of evolution resulting from different curricular and instructional strategies. *Journal of Research in Science Teaching*, 33(8), 879-900.

Jonassen, D.H., Reeves, T.C., Hong, N., Harvey, D., & Peters, K. (1997). Concept mapping as cognitive learning and assessment tools. *Journal of Interactive Learning Research*, 8 (3/4), 289-308.

Kaput, J., Bar-Yam, Y., & Jacobson, M.J. (1999). *Planning documents for a national initiative on complex systems in K-16 education (Report to the National Science Foundation*).

Cambridge, MA: New England Complex Systems Institute. Online source: http://necsci.org/events/cxed16/cxed16.html

Kaput, J., Bar-Yam, Y., Jacobson, M., Jakobsson, E., Lemke, J., & Wilensky, U. (2002). *Two roles for complex systems in education: Mainstream content and means for understanding the educational system itself*. Report on Project #REC-9980241, presented at the Annual Conference of the New England Complex Systems Institute.

Kauffman, S. (1995). *At home in the universe: The search for the laws of self-organization and complexity*. New York, NY: Oxford University Press.

Keil, F.C. (1979). *Semantic and conceptual development*. Cambridge, MA: Harvard University Press.

Klopfer, E., & Colella, V. (2000). *Modeling for understanding*. Paper presented at the Society for Information Technology and Teacher Education (SITE-2000) Conference, San Diego, CA.

Klopfer, E., & Um, T. (2000). Young adventurers: Modeling of complex dynamic systems with elementary and middle-school students. *Proceedings of the Third International Conference of the Learning Sciences* (pp. 214-220).

Klopfer, E., & Um, T. (2002). Young adventurers: Modeling of complex dynamic systems with elementary and middle-school students. *Proceedings of the Fifth International Conference of the Learning Sciences*.

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: a critical appraisal. *Learning and Instruction*, 11, 357-380.

Limón, M, & Carretero, M. (1997). Conceptual change and anomalous data: A case study in the domain of natural sciences. *European Journal of Psychology of Education*, 12(2), 213-230.

Margolis, E. (1998/1999). How to acquire a concept. In Eric Margolis & Stephen Laurence (Eds.), *Concepts: Core readings* (pp.549-568). Cambridge, MA: The MIT Press.

Mayer, R.E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology,* 90(2), 312-320.

McClure, J.R., Bell, P.E. (1990). *Effects of an environmental education-related STS approach instruction on cognitive structures of preservice science teachers*. University Park, PA: Pennsylvania State University. (ERIC Document Reproduction Services No. ED 341 582).

Merriam, S.B. (1998). *Qualitative research and case study applications in education: Revised and expanded form case study research in education*. San Francisco: Jossey-Bass Publishers

Metz, K.E., (1998) Emergent understanding and attribution of randomness: Comparative analysis of the reasoning of primary grade children and undergraduates. *Cognition and Instruction*, 16(3), 285-365.

Miles, M.B. & Huberman, A.M. (1984). *Qualitative data analysis*. Beverly Hills: Sage Publications.

Mosenthal, P. & Kintsch, I. (1992a). Using knowledge modeling as a basis for assessing students' knowledge. *Journal of Reading*, 35(8), 668-678.

Mosenthal, P. & Kintsch, I. (1992b). Understanding knowledge acquisition from a knowledge model perspective. *Journal of Reading*, 35(7), 588-596.

Mosenthal, P.B. & Kirsch, I.S. (1992c). Learning from exposition: Understanding knowledge acquisition from a knowledge model perspective. *Journal of Reading*, 35, 356:370.

Odom, A.L., (1995). Secondary & College Biology Students' Misconceptions About Diffusion & Osmosis. *American Biology Teacher*, 57 (7), p.409-15.

Papert, S. (1980). *Mindstorm: Children, computers and powerful ideas*. New York: Basic Books Inc.

Patton, M. Q., (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage Publication.

Penner, D.E. (2000). Explaining systems: Investigating middle school students' understanding of emergent phenomena. *Journal of Research in Science Teaching*, 37(8), 784-806

Penner, D.E. (2001). Cognition, computers, and synthetic science: Building knowledge and meaning through modeling. In W.G. Secada (Ed.), Review of Research in Education 2000-2001, Volume 25 (pp.1-36). Washington, DC: American Educational Research Association.

Perkins, D.N. & Grotzer, T.A. (2000). *Models and moves: Focusing on dimensions of causal complexity to achieve deeper scientific understanding*. Paper presented at the Annual Conference of the American Educational Research Association, New Orleans, LA.

Pfundt, H., & Duit, R. (1994). *Bibliography. Students' alternative frameworks and science education*. (4th ed.) Kiel, Germany: IPN.

Piaget, J. (1975). *L'équilibration des structures cognitives. Problème central du development.* [The development of thought: Equilibration of cognitive structures]. Paris: PUF (English translation New York: Viking Press).

Pintrich, P. R., Marx, R. W. & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. Review of Educational Research, 63(2), 167-199.

Reiner, M., Slotta, J.D., Chi, M.T.H., Resnick, L. (2000). Naïve Physics Reasoning: A commitment to substance-based conceptions. *Cognition and Instruction*, 18(1), 1-34.

Renner, J.W., Brumby, M., & Sheppherd, D.L. (1981). Why are there no dinosaurs in Oklahoma? *Science Teacher*: 48(9), 22-24.

Resnick, M. (1994). Turtles, termites and traffic jams: Explorations in massively parallel microworlds. Cambridge, MA: MIT Press.

Resnick, M., & Wilensky, U. (1993). Beyond the deterministic, centralized mindsets: New thinking for new sciences. Presented at the annual conference of the American Educational Research Association, Atlanta, GA.

Resnick, M., and Wilensky, U. (1997). Diving into complexity: Developing probabilistic decentralized thinking through role playing activities. *Journal of the Learning Sciences*, 7, 153-172.

Ruiz-Primo, M.A. & Shavelson, R.J., (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569-600.

Rumelhart, D.E. & Norman, D.A. (1976). *Accretion, tuning and restructuring: three modes of learning*. (Report no. 7602). (ERIC Document Reproduction Service No. ED134902)

Sanger, M. J., Brecheisen, D.M., & Hynek, B.M. (2001). Can computer animations affect college biology students' conceptions about diffusion and osmosis? *American Biology Teacher*, 63 (2), p.104-09.

Settlage, J. (1994). Conceptions of natural selection: A snapshot of the sense-making process. *Journal of Research in Science Teaching*, 31(5), 449-457.

Shaughnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Macmillan.

Sinatra, G.M., and Pintrich, P.R. (2002). *Intentional conceptual change*. Las Vegas: Lawrence Erlbaum.

Slotta, J.D. & Chi, M.T.H. (1996). Understanding constraint-based processes: A precursor to conceptual change in physics. In G.W. Cottrell (Ed.), *Proceedings of the 18th annual conference of the Cognitive Science Society* (pp. 306-311).

Slotta, J.D. & Chi, M.T.H. (1999, April). Overcoming robust misconceptions through ontological training. Paper presented at American Educational Research Association annual meeting in Montreal.

Slotta, J.D., Chi, M.T.H., & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition and Instruction*, 13(3), 373-400.

Sommers, F. (1963). Types of ontology. *Philosophical Review*. 72, 327-363.

Spiro, R.J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In *Tenth annual conference of the Cognitive Science Society* (pp. 375-383). Hillsdale, NJ: Lawrence Erlbaum Associates.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1992). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In T. M. Duffy, & D. H. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 57-75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Stieff, M. & Wilensky, U. (2002). *ChemLogo: An emergent modeling environment for teaching and learning chemistry*. Paper presentation at fifth biannual International Conference of the Learning Sciences, Seattle, WA.

Strike, K. A. & Posner, G. J. (1985). Conceptual change view of learning and understanding. In L. West & A. L. Pines (Eds.), *Cognitive structures and conceptual change* (pp.211-231). Orlando, FL: Academic Press, Inc.

Strike, K. A. & Posner, G. J. (1992). *A revisionist theory of conceptual change*. In R. Duschl, & R. Hamilton (Eds.), Philosophy of Science, Cognitive Science, and Educational Theory and Practice. Albany, NY: Sunny Press.

Suits, J. P., (2000). *Conceptual change and chemistry achievement: A two-dimensional model*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Tao & Gunstone (1999). Conceptual change in science through collaborative learning of the computer. *International Journal of Science Education*, 21(1), 39-57.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4, 45-69.

Vosniadou, S. & Brewer, W.F. (1994). Mental models of the day/night cycle. *Cognitive Science*, 18(1), 123-183.

Vosniadou, S. & Ioannides, C. (1998). From conceptual development to science education: A psychological point of view. *International Journal of Science Education*, 20(10), 1213-1230.

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papdemetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction*, 11, 381-419.

Waldrop, M.M. (1993). *Complexity: The emerging science at the edge of order and chaos*. New York, NY: Touchstone.

White, B.Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction*, 10, 1-100.

Wilensky, U. (1995). Paradox programming, and learning probability. *Journal of Mathematical Behavior*, 14(2), 231-280.

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational studies in mathematics*, 33(2), 171-202.

Wilensky, U. (1999). GasLab: An extensible modeling toolkit for connecting micro- and marcro-properties of gases. In W. Feurzeig & N, Roberts (Eds.), *Modeling and simulation in science and mathematics education* (pp.151-178). New York: Springer-Verlag.

Wilensky, U. (2001). *Emergent entities and emergent processes: Constructing emergence through multi-agent programming*. Paper presented at annual conference of American Educational Research Association, Seattle, WA.

Wilensky, U., & Reisman, K. (in press). Thinking like a wolf, a sheep, or a firefly: Learning biology through constructing and testing computational theories. *Cognition and Instruction*.

Wilensky, U. & Resnick, M. (1995). *New thinking for new sciences: Constructionist approaches for exploring complexity*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Wilensky, U. & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, 8(1), 3-19.

Wilensky, U. & Stroup, W. (2000). Networked gridlock: Students enacting complex dynamic phenomena with the HubNet architecture. *Proceedings of the Fourth International Conference of the Learning Sciences*: Ann Arbor, MI.

Wilson, J.M. (1998). Difference in knowledge networks about acids and bases in year-12, undergraduate and postgraduate chemistry students. *Research in Science Education*, 28(4), 429-446.

Wilson, R.A. & Keil, F.C. (2000). The shadows and shallows of explanation. In F.C. Keil & R.A. Wilson (Eds), *Explanation and Cognition*. (pp.87-114). Cambridge, MA: The MIT Press.

Windschitl, M., & Andre, T. (1998). Using computer simulations to enhance conceptual change: The roles of constructivist instruction and student epistemological beliefs. *Journal of Research in Science Teaching*, 35(2), 145-160.

Yin, R. K. (1994). *Case study research: Design and methods*. Thousand Oaks, CA: Sage Publications.

Zaim-Idrissi, K., Desautels, J., & Larochelle, M. (1993). "The map is the territory!" The viewpoints of biology students on the theory of evolution. *The Alberta Journal of Educational Research*, 39(1), 59-72.

APPENDIX A

DATA COLLECTION INSTRUMENTS

BRAIN TEASER QUESTIONS

Name _____________________ **Date___________________**

DIRECTIONS: You are not expected to know the "real" scientific explanations, however, you may have some personal "theories" or understanding about the following phenomena from science articles, novels or movies. Therefore, please answer these questions using your intuition (best guess) or knowledge from informal learning experiences.

1. How would you explain how ants find and collect their food. What rules do you believe they follow? Try to explain using only the space provided. If you don't know, just make a "X" in the space provided.

2. It has been said that a butterfly flapping its wings in Brazil can jiggle the air and thus can help cause a snowstorm in Alaska. Is this possible? If so, how would you explain this phenomena? What type of rules would permit this to occur? If not possible, what rules do you believe would prevent them from occurring? Try to explain using only the space provided. If you don't know, just make a "X" in the space provided.

3. How would you explain the formation of traffic jams? Are there rules that would direct this type of activity? Try to explain using only the space provided. If you don't know, just make a "X" in the space provided.

4. Suppose large deposits of a cancer-curing mineral were discovered on a distant planet. It is too dangerous and costly to send human astronauts to mine the mineral. If thousands of robots were sent. What type of programming would be necessary to ensure that the robots would be able to find the mineral, mine it, take it back to the space ship and then return to their exploration and mining tasks? In other words, what type of rules and strategies should the robots have to follow? Use as much space as required.

Thank you very much for your cooperation.

APPENDIX B

Coding Templates – EMMF answers & CWWM answers.

Table B.1 Prototypical Answers to Butterfly Questions Coded to EFMM Taxonomy

EFMM	Components of coding	*Butterfly* question
Ontological perspective: Emergent	1) Local interactions among agents, 2) leads to the creation of something that exhibits a differential behavior than those of the component agents; 3) this interaction is made possible due to some type of identification (tagging device), 4) and, communication (flows of information and/or resources).	As an individual butterfly's flapping its wings may create a temporary vortex which could potentially interact with other local vortices. If there is a steep temperature inversion present this initial condition could feed energy into the system.
Control of system Initial causes Decentralized	1) The individual agents are independent of each other, yet they all operate under the same rules; 2) the systems organizes itself through the interactions of these independent agents both with each other as well as with the environment.	Prigogines theory of order generation in highly distributed systems. Rules are not involved. Physical laws, generative mechanisms and initial triggering conditions are involved.
Action effects Non-linear	1) Because the system is organized through individual and independent actions, it is possible that one agent's actions can have exponentially significant results.	Like many atmospheric systems that are chaotic, it can be poised on the cusp of an instability, and it can take only a miniscule nudge to push the system into one "basin of attraction" or another. Once the process is initiated the system will tend to slide towards the center of the chosen basin.
Agents' actions Random	1) Agents at the lowest level appear to act in random fashion.	The initial condition is unpredictable. The single butterfly creating a vortex sufficiently powerful, in a location which will set the chain of events into motion cannot be predicted.
Underlying causes Probabilistic	1) The system organizes itself based on the interactions of the agents as described above, therefore the resulting structure is probable. 2) Like other probabilistic processes, larger numbers over longer time periods are more likely to result in the formation of normal distributions.	If a vortex is created is created it is by chance. If it grows it is by chance. If a large number of vortices are created simultaneously it is more likely that one of them may contain sufficient energy and be close enough to a steep temperature inversion to create an amplifying effect.
Systems' Nature Dynamic	1) Once the system, the recurring structure, emerges it exhibits a more stable quality; 2) yet all the component agents have the potential to be replaced by other similar independently operating agents.	The weather system has many different phenomena that create vortices and temperature inversions. As one vortex dies (is dampened) another one is formed. Only when all the elements interact in a certain way do these events grow to a discernable size to be considered a visible weather pattern.

Table B.2 Prototypical Answers to Traffic Questions Coded to EFMM Taxonomy

EFMM	Components of coding	*Traffic* question
Ontological perspective: Emergent	1) Local interactions among agents, 2) leads to the creation of something that exhibits a differential behavior than those of the component agents; 3) this interaction is made possible due to some type of identification (tagging device), 4) and, communication (flows of information and/or resources).	Cars interact by responding to the car directly in front of them. The rules of operation are as simple as respond to a signal, red tail lights. Therefore, when an individual driven sees red brake lights go on they too must put on their brakes. This starts a chain of events in which the drivers behind also respond to this red brake light, an so on and so on. This flow of information from one driver to the next creates a wave of cars with decreasing speeds.
Control of system Initial causes Decentralized	1) The individual agents are independent of each other, yet they all operate under the same rules; 2) the systems organizes itself through the interactions of these independent agents both with each other as well as with the environment.	All drivers must operate under the same rules otherwise there will be not only traffic jams but fatalities as cars crash into each other.
Action effects Non-linear	1) Because the system is organized through individual and independent actions, it is possible that one agent's actions can have exponentially significant results.	If one driver chooses to slow down, as indicated by their bake lights, then all drivers behind them must slow down as well.
Agents' actions Random	1) Agents at the lowest level appear to act in random fashion.	The behavior of the individual driver is totally unpredictable. There is no way to determine ahead of time what of many possible things could make an individual driver slow down. (discomfort, distraction, disregard for rules, external conditions).
Underlying causes Probabilistic	1) The system organizes itself based on the interactions of the agents as described above, therefore the resulting structure is probable. 2) Like other probabilistic processes, larger numbers over longer time periods are more likely to result in the formation of normal distributions.	Once the initial conditions establishing the slowing down of an individual car occurs, the formation of a traffic jam is dependent on many different factors, however, it is never certain that this simple act alone will cause a traffic jam. It may not if the driver resumes speed, or changes lanes, etc. However there are factors which will make it more likely that the initial condition will form into a traffic jam. One of these is numbers. The larger the number of cars on the road, the more likely this initial action will cause a jam. Another is alternative routes available. If there are multiple lanes available it is less likely that the initial condition will result in a jam.
Systems' Nature Dynamic	1) Once the system, the recurring structure, emerges it exhibits a more stable quality; 2) yet all the component agents have the potential to be replaced by other similar independently operating agents.	Cars are always on the road and they are always slowing down and speeding up. Therefore these signals of red bake lights are always going on and off. Therefore to have a single incidence of this slowing down produce a traffic jam will be dependent on a variety of things, one of them is time of day. During certain times day the volume of cars increase therefore the likelihood of forming a jam increases. Once a jam is formed it maintains itself by acting as a backward moving wave: as cars in front leave the jam, cars at the rear enter the jam. When the volume of cars is reduced, the potential for a jam is still there, it is just at an insufficient numbers to reach that critical self-organizing point. Therefore the traffic system exists without the traffic jam.

Table B.3 Prototypical Answers to Robot Questions Coded to EFMM Taxonomy

EFMM	Components of coding	*Robots* question
Ontological perspective: Emergent	1) Local interactions among agents, 2) leads to the creation of something that exhibits a differential behavior than those of the component agents;0 3) this interaction is made possible due to some type of identification (tagging device), 4) and, communication (flows of information and/or resources).	Individual robots would communicate with each other and with the ship through some type of signaling system. The signals would enable them to identify each other, the local of the ship, the discovery of gold deposits, locations that are easier, trails to follow, etc. Once gold is located the signal can be used to draw other robots to the location and form some type of physical trail of robots moving back and forth from the deposit to the ship.
Control of system Initial causes Decentralized	1) The individual agents are independent of each other, yet they all operate under the same rules; 2) the systems organizes itself through the interactions of these independent agents both with each other as well as with the environment.	All robots would be programmed to the same thing: search for the particular markers, if sufficient gold deposit is identified send out signal to draw other robots to the site. In the case of multiple signals respond to the closest one.
Action effects Non-linear	1) Because the system is organized through individual and independent actions, it is possible that one agent's actions can have exponentially significant results.	If one robot identifies the location of a gold deposit, the signal will draw many others to the site.
Agents' actions Random	1) Agents at the lowest level appear to act in random fashion.	The individual agent is programmed to search randomly until they identify a deposit or they are attracted by some signal.
Underlying causes Probabilistic	1) The system organizes itself based on the interactions of the agents as described above, therefore the resulting structure is probable. 2) Like other probabilistic processes, larger numbers over longer time periods are more likely to result in the formation of normal distributions.	Once the initial signal is sent out, the system of multiple robots working together is dependent on the number of robots that are in the location to receive the signal. It is also dependent on the number of other signals that may be within the system (other robots also sending signals). Once sufficient numbers of robots are working together, they will attract more robots to join them by amplifying the initial signal. In addition, as the trail of robots grows the physical obstacles within the environment will become worn thereby making the path easier to move along.
Systems' Nature Dynamic	1) Once the system, the recurring structure, emerges it exhibits a more stable quality; 2) yet all the component agents have the potential to be replaced by other similar independently operating agents.	In this process, an individual robot may at any point be part of the digging crew, transportation crew or delivery crew. Once a site is exhausted, the individual will go back to randomly searching until there is another discovery and the formation of another crew.

Table B.4 Prototypical Answer Coded to CWMM Taxonomy

CWMM	Components of coding	Example for Butterfly question
Ontological perspective: Reductive	1) agents' act in isolation. 2) simple stepwise description.	Storms are local. One to one relationships. Actions are cumulative therefore one butterfly is too small to matter.
Control of system Centralized	1) orders/controls come from outside. Or is within the system but not attributed to the individual agents within. E.g., different agents have different rules. e.g. mention of hierarchy.	Weather systems are controlled by higher level (top down) forces.
Action effects Linear	1) one thing leads to another. E.g. direct link between controller and controllee. e.g., action→reaction	Small actions and small size cannot affect large systems.
Agents' actions Predictable	1) agents' actions are predictable. e.g., they (it) will perform the action. There is no mention of randomness or chance in their actions.	Implication that agents' actions can be calculated and factored out.
Underlying causes Teleologic	1) it knows the end point. E.g., it knows it has to survive.	Underlying cause of storms cannot be attributed to agent levels probabilities. Storms are determined by larger forces outside our control.
Systems' Nature Static	1) explicit descriptions of non changing system.	The effects of the butterflies are local and therefore do not account for changes to the system. All actions are local and terminate.

APPENDIX C

Procedure Used in Constructing Coding Taxonomies

We collected data from both pre and post intervention interviews. All interventions were audio and videotaped in order to permit the researcher to closely observe the interactions and reactions of the subjects to the intervention. Several written documents were also produced by the subjects as another data set in the qualitative case study: (1) delayed posttest results from emergent framework questions; (2) final posttest results from emergent framework questions including one on evolution; (3) concept maps of complex systems concepts; and (4) student records including Nelson-Denny Reading test, and course grades. The following description of the development of the two instruments used to assess the students explanatory frameworks (EFMM, CWMM) and the CST are taken from the first author's dissertation (Charles, 2003).

As the principle researcher I began the data analysis process by first looking at my purpose statement and made the decision to focus on evidence of mental models (e.g., observations, explanations, vocabulary use, analogies, and relationships of concepts). This should not suggest that I ignored the possibility of emerging data categories that more aptly describe the evidence. For instances, categories related to epistemological beliefs, need for social interaction, and coaching behavior. Although these categories are interesting, they lie outside of the scope of this report. However, they may provide new directions for future studies.

The Process of Constructing Categories

The process of constructing categories from the raw data started with the following procedure: transcripts from one student were annotated and a preliminary coding scheme was recorded in a coding logbook. Because of the theoretical nature of the research design, *a priori* coding schema were used to develop certain categories. One such category used the complex systems taxonomy (CST). The decision to use this taxonomy, rather than the ontological mental models taxonomy (EFMM and CWMM), was made because it provided a broader palette from which to describe emergent causal processes. Another *a priori* category was cognitive strategies. In the first round of coding they were identified as 'descriptions' and 'sense-making'. Also identified were the categories of emotional response, for example, frustration and fatigue, as well as the larger category defined by the coaching itself. Development of a coding scheme for the latter category was put on hold until the categories for the learners were fully developed.

Undertaking a second round of coding using the same documents provided a finer articulation of the cognitive strategies category. 'Sense making' was elaborated into concepts of 'paraphrasing', 'explanation', and 'analogies'. I then went back to the literature to look for further theoretical descriptions and explanations of these cognitive processes. Entwistle (1988), and Marton (1981) provided insight into the cognitive processes involved in concept formation, whereas Keil and Wilson (2000) provided me with greater insight on possible ways to code for the category heading of 'explanation'.

The category of 'emotional responses' was changed to 'social interaction behaviors' and 'motivation' (motivated to participate because of: need for social contact; feelings of importance; money; interest in topic) and defined to include such concepts as 'anxiety level', 'tolerance for ambiguity', 'need to please', 'feelings of contribution', 'need to appear smart'. In this round of coding, another category appeared to emerge, that of epistemological beliefs.

The third round of coding took the categories developed in the first and second round and applied them to two other case studies in the cohort. The two selected were believed to be quite different from the original case. The data fit the categories, as defined, and few data points remained uncategorized. Nonetheless, there was a further articulation of the cognitive engagement category where it was felt that metacognitive strategies were being used in the sense-making process. Furthermore, there appeared to be examples of what could be described as "meta-model" thinking[13] where the student attempted to use the new explanatory framework to problem solve using the computer models representations as analogies. It was therefore decided to make this a category unto itself. A testable coding scheme appeared to emerge (Tables 3, 4, and 5) through these repeated cycles of testing and refining or modifying or eliminating the categories.

Category reliability check. After the third round development I decided to conduct a reliability check on the categories. Selecting a totally new transcript (i.e., one that was not part of the database used for the study and not transcribed by the research assistant), I met with the research assistant (RA) and provided him with some background information concerning the types of things that had been coded; for example, evidence of cognitive strategies such as explaining through examples, or evidence of complex system's concepts. The RA was asked to use his judgment and intuitions to identify any other categories that he recognized from the

[13] I suggest that this metacognitive activity is distinct from metaconceptual awareness since the focus is on thinking or reasoning with the representational model rather than evaluating the explanatory framework.

transcript. In other words, to let the categories and coding emerge from the data itself - a grounded theory approach.

What followed was a three-stage process. First, we each independently read and annotated our copy of the same seven-page transcriptions. Then I looked at the coded document and discussed his coding decisions. The RA had made several interesting and unique observations, which were discussed and evaluated based on their significance to the generation of useable categories. The categories that appeared to be most fruitful were added to the already existing list of categories. Some categories such as psychological interpretations were left out because of their subjective nature and the belief that they are not central to the research.

The final task was to compare the documents coded by the RA. Although the RA did not address all the complex systems categories, the agreement on the other categories was very high. In several instances we discussed the name assigned to the coded passages or words that describe sub-elements of the major cognitive strategies category, in other words, the dimensions of the category. By the end, there was consensus on the segments of text that were coded and the category assignment of those segments of text.

The Process of Development Themes

The process of developing themes was informed by Merriam (1998), who tells us that the importance of themes is to test out explanations and hypotheses through the linking of categories. It is important to note that as the themes emerged they also influenced the types of questions that could be explained by the data, hence modifying the central questions of the qualitative phase of this dissertation study. Table 14 (also referred to as a "data display" in Anfara, Brown, & Mangione, 2002) demonstrates another way that I explored the potential hypotheses and explanations as I attempted to link the categories of data. This also was a way to ensure having multiple data sources for triangulation of the data analyses.

Table 14. Finding themes within the data and testing of possible links.

Questions that could be explored through the linking of data.	Pretest/ Posttest	Transcript	Concept Map	Delayed Posttest Study 2	Nelson Denny	GPA	NYA science courses
1. Interaction of sessions and learning of Emergent Causal Processes (ECP) affordances of simulations		X	X				

2. Learning of ECP and transfer to explanatory framework, elaborations of EFMMs		X	X	X			
3. Interaction of students' profiles and ECP.		X	X	X	X	X	X
4. Use of analogous models and improved understanding of ECP.		X		X			
5. Affordances of understanding specific EFMMs concepts - emergent self-organ. - probabilistic causes - dynamic nature - decentralized control - random actions - nonlinear effects	X			X			
6. Synthetic mental models.	X	X	X	X			
7. Correlations of coding schema OMMT and CST.		X		X			

Verification of data. The protocol for verifying the data collected from each instructional session was as follows: (1) meet with non-participant observer at the end of each week to discuss the progress of the sessions; (2) critically review field notes and make a random check of audio and video quality; and (3) use the reflections from the field notes and discussion with non-participant observer to make modifications for the next session.

Confidentiality of data. The researcher upheld the requisite "human research ethics" measures to insure confidentiality of the entire data corpus. All students are referred to by their pseudonym and raw data were shared only with my committee members and coders.

Ethics of conducting human research. Because my role was that of a researcher /coach and not a teacher, I did not actively insist that the student attend to faulty mental models, however, I prompted for reconsideration of their statements (e.g., "do you really think that is what would happen?"; "could you explain that to me again."). However, in my commitment to do no harm, if a student appeared to be creating a new mental model (e.g., "aha, I never thought it was that way…"), which was a classic misconception, I would actively intervene to halt this construction before it was reinforced.

Other considerations. As a researcher with a limited science background, it was important to ensure the veracity and accuracy of the science content. All the instructional materials pertaining to the science content were authenticated and accepted by three members of the science program faculty (i.e., one from each of the major departments: biology, chemistry and physics). They also evaluated the credibility of the selected resources (i.e., web-sites, handouts, etc.). In addition, these same individuals, along with an expert in the field of cybernetics were responsible for the group instruction delivered in the first phase of the research.

During the case study sessions, since I was the sole administrator, I curtailed discussion that pertained to direct science phenomena that reached beyond my level of accurate knowledge. I also was cautious with my use of scientific analogies. Whenever, questions or discussion moved into areas of science, or complex systems, that were beyond my scope, my response was generally "to get back to them with an answer". Fortunately, these situations did not arise very often, and when it did, it was primarily with one student whose cognitive skills were far in advance of the norm. I have identified only one occasion where I misunderstood this student's questioning and mistakenly provided him with an inaccurate answer, which I attempted to rectify in our subsequent session.

Attempting to reduce the first and third threats to validity, this study used multiple measures to generate the data corpus over an extended period of time. Because the data collection process spanned almost two years it allowed me to collect a substantial amount of evidence of the students' changes over time. I was able to establish a good rapport with the nine students thereby increasing the "emic" component of the data. Subsequently gathering evidence of both confirming and disconfirming cases.

The second recommendation to improve validity was addressed by having several different individuals code the data to obtain an inter-rater reliability correlation score. Lastly, accuracy of the facts and interpretation of the transcripts was attempted by having the participants themselves review the data – what Creswell (1994) calls "members checks". Because the final session occurred months after the case study intervention, the written transcripts were available for the students to review. Additionally, I provided the student with his/her summary case reports and reviewed it with each respectively. This allowed for some corrections as well as confirmation of their opinions as to how they had experienced the intervention. A final effort to confirm conclusions regarding concept maps was made by sending JPEG versions to the student by email. I requested feedback if the maps were not representative of their current understanding. No one requested changes, although I did receive a few emails of salutations confirming that they had received the mailing.

APPENDIX D

Examples of Metacognitive Discourse Representing Cognitive Struggle

Example from Greg. During session 2 he articulated his beliefs in a running discourse that took him through the untangling of where the element of randomness was attributable and where it was not.

G: No, it’s like making, um, axioms or something. if you start off with something like although you say it’s the probability of it happening, it will happen over time. It’s not really a probability as much as it depends as a rule.

I: But, isn’t the rule being given to the individual?

G: But even though the rule is being given to the individual. I guess like, lets say if I said, if I switch the nose angle up to 180. Its increases the probability of them conglomerating into groups…

I: Mm-hm.
G: With them staying there. But the thing is for the individual it increases that but the whole system, it means it will have those groups, like what are the chances.
I mean it just make it kind of a rule, like I switch it and it happens. It’s not really the chance of it, it has an increased chance but that increased chance makes it a rule. It’s like an axiom of it.

I: OK. Are you suggesting that within this particular system it becomes deterministic?

G: Yeah. If you change certain things.

I: OK.

G: For the individual it’s still chance, it alters your chances for the whole system

I: You think so?

G: I find, well look at this, here I have a nose angle like that … see at zero how quickly it expands, right away.

I: OK. Let me understand this more clearly.

G: It’s because the individual, like because the system has order. So like order isn’t based on chance, order is based on rules, more … So by changing the chance of the system, no, by switching the chance of the individual you change the order of the system

I: OK.

G: Which is why, like last class [session] I said that these things like they’re not really affecting the individual as much as they are affecting the system because they change the order of the system whereas they only change the chance of the individual. Like chance isn’t always the outcome, it’s just the most likely outcome.

I: So, if I’m understanding you correctly, if we’re looking at the different levels, at the level of the individual, it’s probability…
G: yeah.

I: But on the level of the system , it almost becomes deterministic?

G: yeah.

G: yeah. And that’s the whole basis of a complex system though, isn’t it? Like it has um, like the individuals wont mirror the system itself. The system itself has a sense of order. It’s not determined by the individual.

His understanding of the concept of randomness again showed signs of challenge and development in session 5 when constructing his concept map. At this point he recommended removing the term “random” since he didn’t see it as truly representative of complex systems’ behaviors. He decided that the term “unpredictable” should replace it. This level of debate was a very sophisticated one. As described under the heading of affordances for learning emergent characteristics, I point to this as an example of the limitations of the Multi-Agent simulations employed in this study. Hence the comprehensive understanding of the concept of randomness may be difficult to acquire because the simulations have not built in the generation random “noise” within the system. Therefore, I did not judge this as an example of Greg’s inability to understand predictable actions versus unpredictable (random action) environments, rather, his coming to terms with the representation presented by this conceptual model; therefore, a challenge for him to overcome.

(looking at his old concept map)

G: Take off random. I really wouldn’t think that they’re that random anymore.

I: No?

G: Alright. Like I don’t think… I’d say more unpredictable. Am I allowed to add words?

I: Yeah, absolutely.

G: I’d also say that [unpredictable is] not being the same as random.

I: To you random means?

G: It’s just like. Um it’s not even totally unpredictable. It’s just that it’s not always predictable. Whereas random means like totally unpredictable. So this is sometimes predictable.

Eight months after the instructional intervention Greg demonstrated his integration of this concept into his EFMM. Using his knowledge from his science courses, particularly biology, he expressed a substantial development in his understanding of the concept of randomness.

G: OK, I think the one thing that I’d do, is that I would add random to the single system, to the simple system.

I: Why?

G: Because uh… There’s the uh, the factor of change involved. And just like the small, minute things that each uh, each simple system does… That, that will get uh, like, absorb into the complex system without really having any real effect on it, unless there’s a lot of random events. But uh, you know there has to be randomness somewhere, it’s not like, as I’m far I’m concerned, I mean random events happen.

I: OK, and what’s made you change this idea?

G: Um, well I don’t know why I took random out in the first place. I really, can’t remember. So uh… So for me it’s not really changing it, it’s just uh… Just you know…

The coach takes out the notes and reminds him of the debates he had concerning this concept. Since he cannot remember exactly what his original thinking was, he was asked to explain once again how he now wished to construct his concept map. Before taking biology Greg did not appreciate the role that random noise, like mutations, played in providing the system with opportunities for variety thereby changing the direction of the system's outcome. In session 6 however, he was able to eloquently express the function of randomness within the complex system stating: "it creates possibilities". He also appeared to have integrated this concept with that of probabilities which provide the system with a sense of order.

G: OK, so you want me to explain it like a friend to you?

I: Yeah. (laughs)

G: OK, well, I'd say, a complex system is uh, is made up of simple systems. Of which it can be uh, I'm trying to get away from this light but not in my face, OK… A complex map is made up of simple systems. And these simple systems are random, and dynamic. Like, they follow simple rules, but there's also the whole probability of chance. Like chance is a factor. And so that creates um, randomness, and that creates possibilities, also. That if there were no random events, then you wouldn't have those possibilities. Um, but all these chance events, they, when they get absorbed into the complex system, they have very little effect. It's like throwing a pebble into a river. Sure, you might course the river in a one in billion chance or something, but chances are it does nothing. It's not going to affect the flow of the river in any way. Uh, so, what that means is that complex systems, they follow more rules of probability, and they, they… They I kind of guess being mathematically defined, with algorithms I guess, because they're more likely to, have, um, a real sense or order, that the simple system itself won't have. So what that means is that it's self-organized, and it's uh, um… It's called emergent properties. Um… A complex system is obviously a system, I think that's, a little bit uh, redundant. And there is the chance of uh, unpredictable events, for example, you did throw the pebble. That pebble might stop the flow of the water, by the grace of God or something, so nothing is for sure I guess, there is always the element of chance involved. But they're by and large more predictable than simple systems. And… I think I used everything here.

In a closing comment regarding this concept Greg says:

G: I guess it's just, they've become morel like, more apparent to me, like… That you need probability and chance, I guess I just finally realized it today a little bit more, that you need, like there has to be, different levels, otherwise it's like it's not a complex system. That's the whole, notion of it. And that there's some type of emergent properties in the system. Because otherwise you wouldn't use a system to, to describe it.

.