A Computational Model of Memetic Evolution: Optimizing Collective Intelligence

Noah Welsh
Clemson University, noahwelsh87@gmail.com

A COMPUTATIONAL MODEL OF MEMETIC EVOLUTION: OPTIMIZING COLLECTIVE INTELLIGENCE

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctorate of Philosophy
Educational Leadership

---

by
Noah H. Welsh
May 2014

Accepted by:
Russ Marion, Committee Co-Chair
Joshua Summers, Committee Co-Chair
Leslie Gonzales
Jon Christiansen

ABSTRACT

The purpose of this study was to create an adaptive agent based simulation modeling the processes of creative collaboration. This model aided in the development of a new evolutionary based framework through which education scholars, academics, and professionals in all disciplines and industries can work to optimize their ability to find creative solutions to complex problems. The basic premise follows that the process of idea exchange, parallels the role sexual reproduction in biological evolution and is essential to society's collective ability to solve complex problems. The study outlined a set of assumptions used to develop a new theory of collective intelligence. These assumptions were then translated into design requirements that were designated as parameters for a computational simulation that utilizes two types of machine learning algorithms. This model was developed, and 200 simulations were run for each of 48 different combinations of four independent variables for a total of 9,600 simulations. Statistical analysis of the data revealed a number of patterns enhancing the simulation agents' collective problem solving abilities. Most notably, agents' collective problem solving abilities were optimized when idea exchange between agents was balanced with individual agent time contemplating new creative strategies. Additionally, the agents' collective problem solving abilities were optimized when simulation constraints did not force the agents to converge upon one potential solution.

DEDICATION

I dedicate this study to my father, an unparalleled teacher and mentor, a true renaissance man, and a Walhalla genius.

ACKNOWLEDGMENTS

I wrote this study to illustrate the power of collective intelligence, a view that suggests I owe credit for any of my intellectual accomplishments to all those who have planted the seeds of ideas in my head. I remember learning calculus as a teenager, a mathematics that revolutionized the way I view the world. I could not have created such an intellectual masterpiece in a thousand life times. Discussing his intellectual accomplishments, Isaac Newton, one of the inventors of calculus, once stated, “If I have seen a little further it is by standing on the shoulders of giants” (personal communication, February 5, 1676). From a certain viewpoint, my work is truly plagiarized--only made possible by the collective efforts of the billions of people that came before me. My work is a microcosm of the phenomenon I have set out to model. I must thank the generations of people who helped spawn the evolutionary ancestors of my work.

As for those I was lucky enough to meet, I will start with my committee. Dr. Marion, I cannot thank you enough for your belief in me and your willingness to provide me with such creative freedom. You are a pioneer in the world of leadership theory, but you are practitioner as well. The educational experience you provided me gives me all the proof I need of the incredible merit of your work. Dr. Summers, you are an intellectual explorer, an inspiration. When I approached you as an Educational Leadership PhD student you embraced the immature, disillusioned, and disorganized teenage engineering student that you had known years before. Dr. Gonzales, you are a world-class teacher. You managed to engage the math kid with Critical Race Theory. You kept me focused on the purpose of my work. Dr. Christiansen, I will call you Jon, because in spite of your adeptness as a teacher and mentor, these relationships will always be secondary. You are one of the greatest friends I have ever had and without question the most loyal man I have ever known.

I would also like to thank my classmates for all their help along the way. I owe countless thanks to Kenyae Reese, Matt Della Salla, Cherese Fine, Ernest Mackins, James Nampushi, James Vines, and especially Lori Pindar, along with numerous others. I owe Dr. Todd May and Dr. Bobby McCormick for continuously challenging my ability to think analytically.

I would like to thank all of my friends for helping me maintain my sanity (I realize you may question your success in accomplishing this task). Matthew Sauer, Exley McCormick, David Sitarski, Michael Brannan, David May, John Isely, Seth Beckley, Cedar Howard, and Catherine Cotrupi--I am forever indebted. I apologize to all of my wonderful friends I have left off.

I need to thank my incredible family. I want to thank my grandfather, Walter Hanstein, the most brilliant and wise man I have ever known. I owe my gratitude to my cousin, Danny Moyle, and his beautiful wife and daughter. Dad, I already dedicated this work to you so I won't say much more. I am glad a mutual enemy brought us even closer together. Mom, I have never known someone so dedicated to their principles. My only complaint with you is that sometimes I think you care too much. I cannot think of a better fault to have. You are a saint. I owe my siblings for teaching me so much. Emily, you are the most sensible person I have ever met. Roy, you are the most passionate.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

The purpose of this study is to model the flow dynamics of ideas as individuals collaborate in attempt to solve complex problems, using a variety of ideas taken from the field of evolutionary biology (Dawkins, 1978; Ridley, 2010). This research will help educators and education policy makers by giving them a framework with which to understand how ideas develop in society and will illustrate the conditions that optimize the potential of the creative process. In their discussion of adaptive systems such as the model employed in this study, Miller and Page (2007) stated, "while there is no imperative for adaptive systems to result in optimal structures, there are likely to be conditions under which simple adaptive systems produce optimal conditions" (p. 241). The study was designed to help uncover those ideal conditions.

The intent of this research is to create an adaptive simulation modeling the processes of creative collaboration and, in turn, to manipulate the inputs in order to find the natural conditions that optimize this creative evolution. The model can then help to develop a new framework through which education scholars, academics, and professionals in all disciplines and industries can work to optimize their ability to find creative solutions to complex problems. Miller and Page (2007) explained the value of this approach stating,

> Even as purely abstract objects, computational models are useful. They provide an "artificial" reality in which researchers can experience new worlds in new ways.

Such experiences excite the mind and lead to the development of novel and interesting ideas that result in new scientific advances (p. 76).

1.1 Primary Research Question

What conditions optimize the rate at which networks of individuals collectively solve problems?

1.2 Research Sub-Questions

1. How should collaborative idea exchange be balanced with individual time contemplating new ideas, in order for both individuals and groups to optimally generate creative solutions to complex problems?
2. Should agents copy problem solutions from agents with the current best solutions or copy solutions from random agents in order to optimize the system's ability to collectively solve problems?
3. What percentage of an old idea should an agent change when individually working on problem solutions in order to optimize a system's ability to collectively solve problems?
4. What percentage of another agent's idea should an agent copy when collaborating in order to optimize the system's ability to collectively solve problems?

1.3 Definition of Terms

Due to the interdisciplinary nature of the study a list of terms has been provided in order to reduce confusion created by the use of technical and philosophical jargon. Definitions do not reflect the only definitions of the terms listed below. Rather, definitions reflect the terms as they are used in this study.

1. **Agent:** A single autonomous component used in an agent based model (Page & Miller, 2007).
2. **Agent Based Model:** A class of computational models that simulates the interactions of multiple independently acting autonomous parts (Page & Miller, 2007).
3. **Artificial Intelligence:** A set of computer algorithms and programming techniques which is intended to create machines capable of solving problems.
4. **Artificial Neural Networks:** A type of artificial intelligence algorithm that is based upon the structure of the human brain, which iteratively builds connections between nodes representing ideas and other phenomena (Haykin, 1994).
5. **Asexual Reproduction**: The biological process for procreation in which an individual creates a perfect copy of itself.
6. **Bonferroni Correction:** A statistical tool used to reduce the rate of type I error when a data set with multiple independent variables is analyzed using multiple null hypotheses (Statsoft Inc., 2013).
7. **Black Box:** A process with clear inputs and outputs that utilizes intermediate mechanisms which cannot be viewed (Page & Miller, 2007).
8. **Chaos:** The phenomenon where components of a system interact to create disorder. These systems are highly sensitive to initial conditions or described with traditional linear models (Holland, 1974).
9. **Collective Intelligence:** The combined cognitive and creative ability of a system of individuals working together to solve problems (Levy, 1997).

10. **Complexity:** The phenomenon where components of a system interact to create new patterns. These systems are extremely sensitive to initial conditions and cannot be predicted or described with traditional linear models (Holland, 1975).
11. **Constructivism**: The epistemological view that individuals construct knowledge out of the social context through which they see the world (Marion, 2012).
12. **Design Guideline:** A loose recommendation to be considered when designing a system.
13. **Design Requirement:** A mandatory standard that must be met when designing a system.
14. **Emergence:** The creation of a phenomenon that is produced by the physical structure of our complex and often chaotic Universe, yet cannot be explained through simple reduction to the properties of its components (Goldstein, 1999).
15. **Epistemology:** The study of how humans acquire knowledge.
16. **Eukaryotes**: Organisms that evolved from prokaryotes that have a membrane bound nucleus and are sometimes multi-cellular.
17. **Evolutionary Psychology:** A sociological perspective that tends to view modern human thought patterns and behaviors as those which gave human ancestors the best chance of survival and, in turn, were molded by the process of natural selection (see sociobiology).
18. **Externality:** The effects of an economic transaction on parties not involved in the transaction (see spillover).

19. **Feedback Loop:** A system that has outputs that are also inputs, thus changing one input can change the output (which is also an input), which in turn changes the output again and creates a cycle.
20. **Fitness Landscape:** A multidimensional surface representing a set of theoretical possibilities for a real world system, having one dimension for each input and one dimension for the output (the fitness), thus problem solving can be viewed as navigating different combinations of input dimensions in order to find the combination which yields the highest possible output fitness variable (Miller & Page, 2007).
21. **Genetic Algorithm:** An iterative process that attempts to solve a problem by using trial and error to gradually change a set of multiple combinations of randomly assigned inputs to generate the optimal problem solution. Unlike a hill climb algorithm that only iterates on a single solution, this algorithm iterates on multiple solutions allowing for the recombination of inputs between different solutions (Holland, 1975; Mitchell, 1997).
22. **Global optimum:** The point on a fitness landscape with the best possible output value (Miller & Page, 2007).
23. **GNU Octave:** A high-end programming platform that is an open source version of MATLAB (Eaton, 2013).
24. **Heterogeneous Agents:** Agents that operate differently in a simulation (Page & Miller, 2007).

25. **Hill Climb Algorithm:** An iterative process which attempts to solve a problem by using trial and error to gradually change a randomly assigned set of inputs to generate the optimal problem solution.
26. **Homogeneous Agents:** Agents that operate identically in a simulation (Page & Miller, 2007).
27. **Homophily:** The sociological assertion that humans gravitate towards like individuals.
28. **Induction:** The process of acquiring knowledge through iterative processing of sensory data (Popper, 1979).
29. **Laminar:** The flow of a fluid that is moving in smooth parallel layers.
30. **Linear Systems:** A system with behavior that can be predicted by the additive effects of individual components.
31. **Local Optimum:** A point on a fitness landscape with an output value that is not the best on the entire landscape but is better than output values found at neighboring points.
32. **Machine Learning**: A set of evolutionary artificial intelligence algorithms that uses trial and error to find ideal solutions to specifically defined problems (Mitchell, 1997).
33. **Meme:** An idea which replicates and proliferates itself between minds (Dawkins, 1976).
34. **Memeplex:** A systems of ideas, memes, which act mutualistically to maximize the ability of the system and thus the constituent memes to replicate (Speel, 1995).

35. **Mutualism:** A relationship between two organisms that is advantageous to the evolutionary well being of both.
36. **Naturalism:** The methodological view that science provides researchers with tools necessary to investigate social systems if incorporated into the modern complexity science paradigm.
37. **Non-Linear Systems:** A system with behavior that cannot be predicted by the additive effects of individual components. These are generally viewed as more complex than their linear counterparts.
38. **Ockham's Razor:** The mathematical and philosophical assertion that all else being equal, the simplest solution tends to be correct.
39. **Ontology:** The nature of existence.
40. **Optimization:** A process to find the set of inputs into a system that create the most ideal output.
41. **Path dependent world:** A complex system in which the past choices of agents influence their future ability to reach goals (Page & Miller, 2007).
42. **Positivism:** An epistemology that posits that humans obtain knowledge through logical manipulation of sensory data (Popper, 1979).
43. **Post-positivism:** An epistemology that posits that humans obtain knowledge through logical manipulation of sensory data while also acknowledging that human biases may influence the process (Popper, 1979).
44. **Pragmatism:** The view that the function of thought is to help accomplish goals rather than to reflect reality (Diggins, 1994).

45. **Prokaryotes:** Simple single celled organisms that lack a membrane bound nucleus and are the ancestors of eukaryotic organisms.
46. **Pseudo-Random Number Generator:** A computer algorithm designed to produce numbers that are indistinguishable from those produced by a truly random series (Gentle, 2004). (Any mentions of random number generation in this study are actually referring to pseudo-random number generation.)
47. **Reductionism:** The idea that all complex systems are mere sums of their components (Anderson, 1972).
48. **Replicator:** Any entity that naturally creates copies of itself and is therefore subject to undergo evolution (Dawkins, 1976).
49. **Scientific Method:** A traditional method for designing experiments and investigating questions based upon inductive logic.
50. **Sexual Reproduction:** The biological process for procreation in which two individuals combine portions of their own genetic material in order to create a new unique organism.
51. **Sociobiology:** A sociological perspective that tends to view modern human thought patterns and behaviors as those which gave human ancestors the best chance of survival and, in turn, were molded by the process of natural selection (see evolutionary psychology).
52. **Spillover:** The effects of an economic transaction on parties not involved in the transaction (see externalities).
53. **Symbiosis:** Interactions between two biological species.

54. **Turbulent:** The flow of a fluid that is moving chaotically and not in smooth parallel layers.

1.4 Overview

The creative, collaborative networks under investigation include graduate education research teams, academic departments, think tanks, and engineering firms. As previously stated, the purpose of the study is to investigate the flow dynamics of ideas in collaborative systems on the cutting edge of specific research questions. Holland (2004) explained that the concept of flow is not limited to the movement of physical fluids, but,

> we speak of the flow of goods into a city or the flow of capital between countries. In more sophisticated contexts, we think of flows over a network of nodes and connectors. The nodes may be factories, and the connectors transport routes for the flow of goods between factories (p. 23).

In the context of education and idea exchange, we speak of the flow of information and creative problem solutions between different people (nodes) along all lines of communication between them (connectors). The reason these ideas are more applicable to graduate education than undergraduate and P-12 education is that the goal of this study is to investigate the creative processes of individuals who are considered to be experts in their given fields. When examining lower levels of education, the majority of ideas trickle down from the higher levels that are actually making these discoveries (Rogers, 2003). Naturally, it can be assumed that most high school geometry classes are not creating genuinely new mathematical conjectures. Rather, they are simply learning the ideas of Euclid. However, in the realm of high level academic research where researchers search for solutions to explicit problems, the quality of the solution becomes much more important in determining what ideas gain traction in the community of specialized experts. By utilizing the ideas of Darwin as well as modern evolutionary biologists, the hope is to model the flow dynamics of the systems that seek to come up with these genuinely novel ideas. By providing this explanatory model, it is possible to provide insight into the conditions that optimize this desired creativity.

This research is intended to help illuminate the phenomenon of collective intelligence. Levy (1997) defined collective intelligence as the cognitive capacity of a group of individuals whom are collaborating and competing, when examined as a whole. Schut (2010) explained that in spite of the diversity of approaches that have been used to address how collective intelligence emerges, a succinct explanation of the guiding mechanisms has yet to be provided. In this context, emergence is defined as the creation of a phenomenon that is produced by the physical structure of our complex and often chaotic Universe, yet cannot be explained through simple reduction to the properties of its components (Goldstein, 1999).

1.5 Theoretical Framework: A Naturalist View of Complexity

This computational model is constructed upon a naturalist ontology. The naturalist position suggests that science provides researchers with tools necessary to investigate social systems if incorporated into the modern complexity science paradigm. Science may not give us the tool set to predict human behaviors. However, the complexity perspective illuminates how independent system components following these mechanical principles interact to form complex entities (Miller & Page, 2007). Unlike complexity based research methods, many traditional research tools assume that systems can be modeled as linear functions of their various attributes (Holland, 1974). Avoiding assumptions of linearity significantly increases the ability of the complexity perspective to describe, model, and predict the behavior of social systems (Marion, 2012). Kauffman (1993) expressed his view of complex systems and their importance stating,

> The overall answer may be that complex systems constructed such that they are poised on the boundary between order and chaos are the ones best able to adapt and mutate by natural selection. Such poised systems appear to be able to coordinate complex, flexible behavior and best able to respond to changes in their environment. I suggest that selection does achieve and maintain such poised systems. Further, beyond the selective molding of individual adaptive systems, there are provocative, promising indications that linked co-evolving complex systems are led by selection, as though by an invisible hand, to form ecosystems whose members mutually attain the edge of chaos. Here all may sustain the highest expected fitness, even while avalanches of co-evolutionary changes propagate through the ecosystem, ringing out old species and ringing in new ones (p. 29).

All social entities, especially those working collaboratively to solve problems fit into Kaufman's description of complex systems. Kaufman (1995) stated that while people may never be able to fully predict the outcomes of complex systems people can discover laws governing their behavior. These complex social systems can be better understood by drawing an analogy to meteorologists attempting to predict the weather. These scientists have a quite sophisticated understanding of particle mechanics, but any layman watching the weather channel can see that they still struggle at predicting hurricanes, despite these weather systems being ultimately reducible to the axioms of particle mechanics. This weather example is consistent with Anderson's (1972) assertion that reductionism, the stance that all complex entities are sums of their components, does not necessitate that understanding the constituent components will provide the ability to understand the complex behaviors which emerge from them. This helps illuminate why many academic fields predict complex phenomena by using models that ignore the base principles out of which these complexities emerge.

In fluid mechanics courses students learn complicated equations filled with arbitrary coefficients that enable them to predict whether fluids flow in smooth laminar lines or in chaotic turbulence. In truth, these numbers are arbitrary and have nothing to do with the fundamental axioms of physics. The students use a model grounded in assumptions of complexity rather than particle mechanics. Even if scientists do uncover a theory of everything, the predicted theory that is supposed to unify the world of theoretical physics, they still will not be able to perfectly predict the weather. They will not be able to predict these complex and chaotic meteorological systems because these systems are extremely dependent upon initial conditions. The task of predicting weather utilizing particle mechanics would require a practically omniscient knowledge of the initial conditions in addition to practically infinite computational capabilities. Complexity science provides the tools necessary to bridge the gap between the fundamental mechanics of individual system components and the intricate patterns and behaviors that emerge from them.

As researchers, educators, engineers, and business managers seek to optimize the creative processes of the social systems in which they work or supervise, they tend to look to traditional statistical approaches. However these models can be problematic. Creative outputs of these social systems generally do not depend on any individual components of the social systems or the problems they are trying to solve. Instead, the creative outputs are produced as a result of complex functions of the interactions between the multitudes of constituent components in the systems at hand. Many statistical models describe systems as being linear in nature, thus yielding these tools limited in their ability to analyze such complexities. As a result of the complex nature of these social systems, many creativity experts are turning to new approaches to model these interactive social systems.

In summary, this research project focuses on the development of a new computational model in hopes to develop our understanding of the ways in which ideas evolve. Due to the fact that this field of study is complex by definition, no single computational model will ever be capable of providing a complete explanation of the phenomena at hand. This research is based on a small set of assumptions about the nature of the Universe, human consciousness, and human collaboration. Boyatzis (2006) showed that individuals working to change organizations are more effective when they have an understanding of the complexity perspective on organizations, in turn, illustrating the value of this study's approach to modeling social systems.

1.6 Organization of Study

This study is composed of five chapters. Chapter 1 provides a brief overview of the study, its purpose, the research questions, research design, limitations, and theoretical grounding. Chapter 2 explores relevant literature in the field of education as well as explores literature on positivism, pragmatism, and various other epistemologies in order to layout the philosophic groundwork for the study. From here, it examines literature drawing parallels between evolutionary biology and the creative evolution of ideas in society, largely emphasizing meme theory (Dawkins, 1968; Blackmore, 1999). Finally, Chapter 2 analyzes literature examining the phenomenon of collective intelligence.

Chapter 3 explains the specific method that this study employed. First, it explains the general principles of computer simulations, agent based models, and machine algorithms. Next, it outlines a set of assumptions the model makes about the nature of collaborative systems and translates them into design requirements, standards which the model must meet in order to be considered valid. This step helps to draw connections between the theoretical positions laid out in Chapter 2 and the research design laid out in Chapter 3. Chapter 3 then proceeds to explain the specific mathematical, computational, and algorithmic mechanics of the research design in order to illustrate how the model meets the design requirements described above. Finally, Chapter 3 explains the statistical tools used to analyze the data.

Chapter 4 describes the data that the simulation produced and proceeds to describe the results of the statistical analysis of this data. Chapter 5 outlines a variety of conclusions drawn from the simulations and the following statistical analyses and frames them within an appropriate discussion of the study's limitations. Here, it proceeds to explain a variety of future research methods, both empirical and simulation based, which could be used to further explore, expand upon, and validate the findings and conclusions of this study.

1.7 Limitations of Study

Simulation based research has a variety of limitations. Computer models help explain the real world by making assumptions about the subjects at hand and use a variety of algorithmic and mathematical tools to explore these ideas. The problem with this approach is that these findings are not empirical. For example, a simulation designed to explain the influence of secondary school leadership structures on student achievement may provide profound insight into the relationships between these variables, but it cannot be considered hard data. Data can only provide conclusive evidence on these relationships if derived from statistically significant patterns in properly controlled empirical research.

The function of simulations is to provide theoretical insights into the behaviors of systems with better efficiency than many traditional research approaches. These simulation results can then be analyzed in the context of the existing body of literature and new theoretical suggestions can be made. These theoretical positions can then be tested through empirical methods. Some agent based simulations have limited precision, but Miller and Page (2007) argued, that the benefit of the extreme flexibility of these computational tools is well worth this cost.

Despite the limitations of simulation based modeling, this approach to research has a variety of advantages justifying its use. In a complex Universe it is mathematically, practically, and economically unreasonable to think we can focus all of our mental efforts into running empirical tests. Hypothetically speaking, it is possible that humans could invest infinite resources continuously for an infinite period of time into empirical research and still not complete all of the needed empirical work necessary to completely understand our Universe. This conclusion can be drawn from Gödel's incompleteness theorem (1962) illustrating that a set of infinity numbers does not necessarily contain all numbers. Instead, there may be infinitely more numbers outside of the set than within. Likewise, it follows that a hypothetical research plan containing an infinite number of studies will not necessarily contain all of the needed studies, and in fact there may be infinitely more needed studies outside of the research plan as inside of it. For this reason, attempting to empirically solve problems through brute force is inadvisable. By implementing simulation studies to search for the proper phenomena to empirically explore, researchers can significantly increase their efficiency in terms of resource and time allocation. Computational simulations that only approximate the conditions of real world systems are often the most inexpensive and, therefore, the most ideal research tools in many circumstances (Miller & Page, 2007).

CHAPTER 2

LITERATURE REVIEW

While computational modeling of creativity is becoming rather prevalent in academia, it is almost unheard of in the field of education. Traditionally, creativity researchers have assumed that the creativity of a system is primarily influenced by the creativity of the individuals by which the system is composed. Christiansen and Varnes (2008) illuminated some of the holes in this view by examining innovation through a network process and showing that ideas do not arise linearly as traditionally conceived. These findings employed Newman's (2003) complexity perspective on social networks and showed that a system's ability to innovate cannot be modeled as a simple direct function of the aptitude of the agents that compose it.

Cheng and Van de Ven's (1996) analysis of biomedical innovations supports these findings as well. Here, they argued that analysis of these innovations reflects a chaotic system in which ideas emerge from nonlinear dynamical systems that were neither fully organized nor completely random and stochastic. These findings help to question the default stance on creativity. The traditional line of thought necessitates that the ability of think tanks, academic departments, and engineering firms to generate new ideas is largely determined by their hiring practices--by their ability to bring in creative people. Uhl-Bien, Marion, and McKelvey (2007) argued that, "Complexity science suggests a different paradigm for leadership--one that frames leadership as a complex interactive dynamic from which adaptive outcomes (e.g., learning, innovation, and adaptability) emerge" (p. 299).

Homophily provides a clear example of how abstract concepts pertaining to social systems may influence the functionality of organizations of people. McPherson, Smith-Lovin, and Cook (2001) described homophily as the tendency of individuals to associate and develop connections with similar others. This suggests that individuals naturally gravitate towards those with whom they share common perspectives. McPherson, Smith-Lovin, and Cook then proceeded to explain this phenomenon in greater detail stating,

> Homophily limits people's social worlds in a way that has powerful implications for the information they receive, the attitudes they form, and the interactions they experience. Homophily in race and ethnicity creates the strongest divides in our personal environments, with age, religion, education, occupation, and gender following in roughly that order. Geographic propinquity, families, organizations, and isomorphic positions in social systems all create contexts in which homophilous relations form. Ties between non-similar individuals also dissolve at a higher rate, which sets the stage for the formation of niches (localized positions) within social space (McPherson, Smith-Lovin, & Cook, 2001).

This in turn limits the diversity of inputs individuals receive in their day-to-day social interactions, thus can limit their capacity for innovation. Marion (2012) stated that homophily tends to prevent the heterogeneous interactions that, "encourage exchanges by introducing interesting topics to discuss, (p. 464)" and that these heterogeneous interactions, "create conflicting constraints thus pressuring agents to adapt to one another's preferences. We found, consistent with hypothesized expectations, that diffusion is optimized when combined vision and interdependency levels are moderated."

(p. 464). Conversely, Rogers (2003) argued that increasing the heterogeneity of groups can make it difficult for ideas to diffuse within them.

Lattuca (2002) emphasized the importance of institutional structures that provide the necessary conditions to destroy barriers preventing interdisciplinary research as well as to generally incentivize the interdisciplinary idea exchange that this model describes. A variety of research has shown the need for interdisciplinary and transdisciplinary collaboration between academic fields in an attempt to break down the barriers that prevent the free flow of ideas. These academic fields often become completely isolated fields with virtually no communication with other academic communities despite similarity of research interests between these disciplines (Austin, 1990; Boyer, 1990, 1997; Clark, 1983; Damrosch, 1995; Dill, 1991; Kerr, 1982; Tierney & Rhoads, 1994). Interdisciplinary collaborations attempt to allow different academic disciplines to work together while maintaining their disciplinary paradigms, whereas, transdisciplinary collaborations attempt to take this process a step further by formulating new sets of paradigms that are logically consistent with all of the collaborating fields. The question remains, what is the value of these cross-disciplinary approaches to idea generation?

2.1 A Pragmatic Approach

A variety of different philosophic positions generated over the centuries describe the ways in which people acquire knowledge. These theories of knowledge acquisition are called epistemologies. Currently, there is a debate in the field of leadership theory regarding creativity between camps holding entity and collectivist perspectives. Marion (2012) argued for the collectivist approach to creativity studies; this assumes new ideas or knowledge, emerge from the complex interactions of social structures rather than out of the independent cognitive processes of individual agents. Marion's (2012) assertion rejects the notion that individuals are the sole generators of new ideas and new knowledge, but this does not yield it mutually exclusive with other epistemologies such as positivism and post-positivism.

The convoluted relationships and shared ground that exists between these epistemologies is perhaps best illustrated by Popper's stance. Popper (1979) accepted the fundamental role of inductive reasoning, a position typically associated with positivism, but denied that humans are capable of employing this iterative form of logic independently from the influence of their social contexts, biases, and subjectivities (sometimes Popper is referred to as a post-positivist). More confusion appears in the discussion as it becomes extremely difficult to identify specific causal relationships between the actions of individual system components and the complex outcomes that emerge from them (Miller & Page, 2007).

In order to avoid confusion and the complications that come with selecting an epistemological label that is already loaded with years of connotations and subtle differences in interpretations, this study will assume a pragmatic position. Pragmatism is the view that the primary goal of thought is to solve real world problems (Diggins, 1994). In accepting this pragmatic position, the remainder of this chapter serves to explain the theory necessary to successfully ground an effective model of collective intelligence.

The goal of this study is to develop a tool that can be used to help people optimize the effectiveness with which collaborative groups solve problems, a goal that the pragmatic epistemology of the study necessitates is more important than any underlying theory of knowledge acquisition. After all, the function of pragmatism is to solve real world problems rather than to focus on trivial academic details. This pragmatic assertion may seem counterintuitive, as the remainder of this chapter is significantly theory laden. However, only the theoretical positions that are inherent to establishing the model's validity are explicated. This model is conducive with various aspects of post-positivism, collectivism, and constructivism, all of which can help to solve the real world problem at hand (the study's pragmatic aim) so there is no need to iron out these subtle differences.

2.2 The "Self" as a Cognitive Agent

Many modern leadership theorists are beginning to shift their focus from models which view creativity as a phenomenon produced by individuals to a new paradigm which attributes creativity to collections of interacting people (Marion, 2012). Ethiraj and Levintha (2011) showed that models which assume overly specific modularity (such as assuming individuals are the loci of innovation) can limit their ability to describe innovation in complex systems. Hargadon and Bechky (2006) illustrated that it is possible to identify occurences in innovative systems where a single individual cannot be identified as the sole creator of an idea. They did this without denying the value of individual contributions to the collaborative process.

By understanding the "self" as a mere label for a particular cognitive agent (Blackmore, 1999), as consistent with the naturalist framework of this study, new possibilities open up in the realm of innovation research which lie beyond the traditional entity based approach to leadership theory. The entity-based approach assumes an overly simplistic model identifying the "self" as the creative locus of new ideas. Here, the "self" may be a useful tool in describing the world, but does not exist in an ontological sense. The "self" or the individual does not exist outside of the constructs of the physical systems from which it emerged; it is a culmination of nature and nurture (Blackmore, 1999). The notion of "self" helps this collection of brain feedback systems navigate the world, but the "self" label that is used to describe the way the world works is just that, a description. Descriptions are useful in their ability to model reality, but have no purpose in deriving further truths of reality. This is true just as classifications of types of educational institutions may help people describe education, but are not a basis for determining the fundamental nature of education.

"Selves" are masses of neurons that run a variety of brain programs. They process inputs and produce outputs. The "self " is not capable of exerting a metaphysically independent will that creatively steers its future course of existence (Dennett, 2004). However, it is perfectly reasonable to assume that this "self", this organized system of organic compounds, does have complex feedback systems which act as the final step in the infinite causal chain (Dennett, 2004). Suppose people are analyzing a situation in which "I" choose to do some action *X*. "I" is simply a label representing one particular brain. Likewise, "choose" is simply a description of the output of a system of complex brain feedback loops. Therefore, when someone says, "I choose" to pursue the action *X*, what they really mean is that this particular organic computer ran a program that set this particular action *X* in motion.

Understanding the “self” as a human construction representing a specific organic computer has profound implications on the way people think about creative enterprises. Destroying the ontological conception of the “self” enables people to realize what the collectivist based approach to creativity research is already showing them. When debating the primary founder of calculus people mention Leibniz and Newton, but fail to mention the thousands of mathematical scholars before them whose ideas they inherited. Once people understand that the idea of the “self” is no more ontologically independent of physical reality than a group of people, or any social structure, they see that choosing the individual as their unit of analysis in creativity research may be practical in some circumstances but is always arbitrary and generally inadvisable.

This philosophic conception that the “self” is nothing more than a label and sociological construction resonates deeply with Latucca’s (2002) view on the positions of sociocultural theorists. She stated that, “by shifting the unit of analysis from the individual toward the sociocultural setting in which learning is embedded, sociocultural theorists train their focus on the structures and interrelations within communities of practice” (p. 712). All creative structures are emergent properties of the complex physical Universe. It is just as logical for people to study the creative processes of small study groups or larger worldwide societal creative collaborations as it is for them to study the creative processes of individuals.

Lattuca (2002) emphasized the need to extend our conception of social and culture phenomena beyond that which focuses on the individual, stating

> traditional psychological theories of learning and studies based on these theories manifest this assumption about the separate spheres of thinking and being. In contrast, sociological and anthropological theories focus intently on contexts and cultures: they are more apt to assume that analytic strategies should begin with an account of social phenomena and then, on the basis of these, develop analyses of individual mental functioning (p.712).

Lattuca's suggestion to shift the sociological unit of analysis from the individual to the group is consistent with Vygotsky's (1981) criticisms of traditional cognitive theory. Here, Vygotsky criticized perspectives that analyze cognitive functions of individuals independently of their social circumstances. Here, rather than rejecting the cognitive conception of the "self", it can be understood as a label describing a particular complex system of neurons which is a small piece of the even more complex, larger collective entity labeled society. These "selves" are the individuals represented by agents in the agent based model that Chapter 3 describes.

2.3 Memes

Organisms evolve by the process of natural selection (Darwin, 1859). Those that effectively reproduce and thus propagate their genes continue to exist in time replacing their less apt competitors. Today, academics often use this line of thought to frame questions. They look at a particular widespread, sociocultural or behavioral pattern, and ask the question: How is this pattern evolutionarily advantageous to the genes of the individuals demonstrating it, and more importantly how was it advantageous to their ancestors? In other words, how does this behavior help individuals survive until they can mate and, in turn, propagate their DNA? This question is certainly reasonable. A human instinctively exhibiting behaviors which increase its odds at DNA proliferation, is no different than a female grizzly instinctively protecting its young in order to better the odds that its genetic sequences will continue to exist.

This field of thought, called evolutionary psychology, has emerged as a popular method for explaining human behaviors. It posits that a large portion of human interactions and social constructs can be explained as behaviors which were produced in the early stages of human development, a period which occupies the majority of human history when people survived as hunter-gatherers. A number of academics have gone on to explain a wide variety of modern behaviors using this analytic approach (Baker, 1996; Buss 1994; Fisher, 1930; Ridley, 1993; Symons, 1979; Trivers, 1972; Wright, 1994).

While evolutionary psychology has done much to explain modern society there is actually a second evolutionary mechanism at work that must be considered. Just like physical traits and genes, behaviors and ideas which naturally replicate themselves are more likely to found in future populations than ones that tend to be forgotten. Berger (2013) illustrated this idea, showing how certain marketing techniques can be used to enhance the virility of an idea, the extent to which it will spread through the masses. In spite of the recent uptake in this new method of social analysis, this line of thinking has existed for over a hundred years.

Individuals adopt the beliefs of other agents on the basis of a few parameters. Agents typically adopt beliefs of their elders and peers through social interaction. Christians are more likely to have Christian children, and Muslims are more likely to have Muslim children. This process which Cavalli-Sforza and Feldman (1981) described as the vertical-transmission of ideas may seem obvious, but has significant ramifications on how people think about the way individuals develop ideas. Cavalli-Sforza and Feldman (1981) contrasted this type of idea propagation with the propagation of ideas between peers. Simply put, individuals adopt the ideas that they adopt. Ideas propagate themselves through the world just any other replicator, like the genetic allele for color blindness through human genes (Dawkins, 1976).

One of the first academic seeds of the field of cultural evolution was planted when Baldwin described learning patterns through the process of imitation as being hereditary in nature (1896, 1909). Tarde (1903) posited that imitation plays a key role in allowing ideas to diffuse through out groups of people, and many others followed to draw out the analogy between this diffusion and biological evolution. Campbell (1960, 1965) argued that cultures evolve independently of the genes of their host humans. The most important contribution to the idea of thought evolution appeared when Dawkins (1976) founded the field of memetics, specifically explaining the basic unit and functional mechanism powering this process. Dawkins (1976) stated,

> What, after all, is so special about genes? The answer is that they are replicators. …The gene, the DNA molecule, happens to be a replicating entity that prevails on our planet. There may be others. If there are, provided certain other conditions are met, they will almost inevitably tend to become the basis of an evolutionary process" (p. 192).

Next, Dawkins explained how ideas are the new replicator on earth, and coined this new replicator the meme. He later proceeded to state:

> Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or building arches. Just as genes propagate themselves by leaping from body to body via sperm or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process, which in the broad sense can be called imitation. If a scientist reads about, or hears a good idea, he passes it on to his colleagues and students. He mentions it in his articles and his lectures. If the idea catches on it can be said to propagate itself. As my colleague N. K. Humphrey neatly summed an earlier draft of this chapter: …'memes should be regarded as living structures, not just metaphorically but technically. When you plant a fertile meme in my mind, you literally parasitize my brain, turning it into a vehicle for the meme's propagation in just the way that a virus may parasitize the genetic mechanism of a host cell.'" (p. 192**)**.

A long standing debate still exists about the nature of the meme, the unit of thought which cannot be reduced any further (Dawkins, 1976; Delius, 1989; Dennett, 1991, 1995; Durham, 1991; Grant, 1990; Hull, 1982; Lynch, 1991; Williams, 1966). It is unlikely that this debate will be resolved until brain scanning technologies develop the resolution necessary to explain exactly how ideas are coded into the mind, but ultimately the propositions of this study are dependent upon recognizing that the meme exists. They do not depend on knowing the specific neurological mechanisms by which memes are stored in the brain.

2.4 The Selfish Meme

Solutions to real world problems are not composed of single components or single memes. The internal combustion engine functions on the basis of a complex system of multiple interacting memes. The nature of the interaction between memes requires examination under the evolutionary paradigm of this study. For a moment return your attention to genetic evolution. In 1976, in the same book which Dawkins introduced the idea of memes, Dawkins proposed the idea of the selfish gene, stating that genes have zero inherent incentive to propagate the existence of the entire genome of their host organism. This can be misleading. No gene has true incentive or intention, as individual genes are not sentient entities. Here, incentive refers to the driving evolutionary force, the mere fact that entities that replicate and preserve themselves are more likely to be found existing in the future. Here, the notion of the selfish gene suggests that the only evolutionary interest of a single gene is to promote its own continued existence. A single genetic nucleotide, the most basic unit of DNA, is only interested in preserving the existence of itself. It has no inherent interest in preserving the welfare of the host organism that carries it around.

While the previous paragraph explained that genes have no inherent interest in promoting their host organism it would be false to say they have no indirect motive to do so. In fact they usually have a very strong reason to do so. Their future success is indirectly dependent upon it. For sake of illustration, consider a single bear is mutated to have a single gene that causes its hair to be white. This single gene has no inherent interest in whether any other genes in the organism survive genetic selection. Suppose the bear wanders up into an area with huge quantities of snow where it then takes a role in producing a population of new white bears this gene has accomplished its evolutionary goal. This is true regardless of whether every other gene in the initial white bear has been eliminated from the population. Now, suppose these new white bears have adapted to feed on seals and the ability to digest seal meat is dependent upon another gene. Although not inherently so, the evolutionary propagation of the white fur gene is indirectly linked to the ability of its host to digest seals and is thus dependent upon the evolutionary success of the seal digestion gene.

Returning to the basic unit of idea exchange, memes follow the same evolutionary self-interest motives as genes. Memes have no inherent interest in the evolutionary welfare of other memes, but the mutualistic relationships they develop do push them to become highly interdependent. Complex systems of mutuallistically interconnected memes are called memeplexes (Speel, 1995). Using meme theory we begin to see the history of mankind unfold in front of us. We realize that the ideas, which enabled Henry Ford's mass production of the automobile, are the logically descended from the log wheels ancient people used to set up Stonehenge, just as much as we are descended from simple prokaryotes.

2.5 How Ideas Evolve

In a letter to Robert Hooke, Isaac Newton once stated, "What Descartes did was a good step. You have added much several ways, and especially in taking the colours of thin plates into philosophical consideration. If I have seen a little further it is by standing on the shoulders of Giants (I. Newton, personal communication, February 5, 1676)."

When thinking about this note in the context of Dawkin's (1976) meme theory, people are likely to picture themselves as agents carrying memes on the current tip of idea chains, chains dating back to the earliest people. As ideas were passed from person to person the creative evolution took place, as some individuals were able to make minor changes to the ideas before passing them on to their successors. The problem with this explanation is that it cannot satisfactorily account for the exponential acceleration in the current development of ideas and technologies.

Journalist and zoologist Matt Ridley (2010) offered a solution to this problem in his book, *The Rational Optimist: How Prosperity Evolves.* In this book, Ridley argued that modern thinkers have significant reason to be optimistic about humanity's future. He attributed the rationale behind this optimism to the sexual reproduction of ideas in modern society that, in turn has lead to the exponential growth of our understanding of the world around us. In evolutionary biology asexually reproducing species evolve much more slowly than their sexually reproducing counterparts. It only took *Homo erectus* a few million year to evolve into *Homo sapiens,* an insignificant amount of time when compared to the billions of year that it took asexually reproducing prokaryotic organisms to evolve into more complex eukaryotic organisms.

Ridley (2010) argued that sex provides organisms with an enhanced ability to produce new combinations of genes, in turn, speeding up the rate of genetic evolution. This advantage of sex can be translated into cultural terms to explain how humanity's progress is accelerating exponentially. Ridley is not the first to suggest that sex plays a vital role in evolution. Charles Darwin's grandfather Erasmus Darwin (1794) stated that, “if vegetables could only have been produced by buds and bulbs, and not by sexual generation, that there would not at this time have existed one thousandth part of their present number of species (p. 519).” Ridley’s accomplishment was to extend these ideas to the social sciences. His argument is that anthropological evolution parallels the patterns of its biological predecessor.

The Acheulean hand axe, a rock used by a particular tribe of ancient people, which had one sharpened end illustrates this phenomenon. This same hand tool was used for half a million years with no improvements. The reason for this phenomenon is that people used this tool before they started exchanging goods and ideas with other cultures (Ridley, 2010). At the point in time, when cultures began to trade and exchange ideas, societies and cultures began to evolve. This cultural transition occurred to people before the invention of farming and is the real transition that enabled people to move beyond hunter-gatherer survival techniques and into modern civilization (Ridley, 2010). Fortunately, this pattern did not stop here, as seen by the telescoping nature of societal progress.

Toth and Schick (1993) ran detailed analyses of a variety of ancient stone, hand tools. After studying the tools they obtained the raw materials needed to make them and spent months experimentally chipping the rocks until they finally discovered a method to make such objects. These results further support Ridley’s (2010) arguments on the difficulty of developing new ideas without the aid of communication and idea exchange. Toth and Schick took months to develop their stone tools even though they engaged in the process with full knowledge of the end product they intended to create, a sizable advantage that the first humans attempting to make such objects did not have. It can be reasonably assumed that this would make it even more difficult for ancient people to create these objects, an assumption which further supports Ridley's (2010) conjecture that idea exchange is paramount to the human ability to solve problems.

Not only has the sexual reproduction of memes by methods of imitation and communication helped humans rapidly increase their ability to solve problems, but it has done it to such an extent that many geneticists and memeticists have argued that it provided such a drastic survival advantage that it pushed humans to both evolve bigger brains and the ability to communicate (Blackmore, 1999). While sign language has been successfully taught to a few apes, these primates do not exhibit the grammatical abilities and comprehension of symbolic logic that even the youngest human children show. Here, humans show unparalleled abilities to communicate and use symbolic logic (Deacon, 1997; Donald, 1991, 1993; Dunbar, 1996; Pinker, 1994; Pinker & Bloom, 1990). These animals are unable to rearrange words in new grammatical structures to express new ideas. They can only be conditioned to imitate basic hang signals.

In addition to the evolution of language, humans have evolved to have a much larger frontal cortex than other animals, the portion of the brain which enables people to simulate outcomes of potential decisions (DeSalle & Tattersall, 2012; Gilbert, 2007). This gives humans the ability to decide what memes to take from their peers by simulating the outcomes of different choices they could make. The introduction of these two abilities to humans increases the influence of the sexual reproduction of memes on human progress even further. This sexual reproduction of ideas provides a basis for helping to develop a new theory of collective intelligence.

2.6 Collective Intelligence

Collective intelligence is a phenomenon that Levy (1997) defined as the cognitive capacity of a group of individuals as they collaborate and compete. It is associated with the collective group's ability to perform a diverse array of tasks (Woolley, Chabris, Pentland, Hashmi, & Malone, 2010) and illustrates many of the attributes of swarm intelligence, the phenomenon where groups of insects organize themselves into complex behavioral patterns which exceed the cognitive limitations of any of the constituent individuals (Bonabeau & Meyer, 2001; Surweicki, 2005). Likewise, humans naturally organize themselves into systems that push their collective intelligence beyond even the smartest individual in the group.

Read (1958) wrote an essay entitled "I, Pencil: My Family Tree as Told to Leonard E. Read," from the perspective of a pencil as it explains the complexity of all of the steps that go into making it. In illustrating the incredible complexity of creating something seemingly as simple as a wooden pencil, Read showed that the creative and cognitive capacity of society at large must be greater than the creative and cognitive capacity of any human on Earth. This must be true since no human could possibly master all the skills and knowledge that Read listed. The this study is to illustrate the memetic mechanisms through which collective intelligence emerges and, in turn, to provide insights into how people might act to optimize their collective ability to solve problems.

Significant empirical evidence exists illustrating the connection between social interaction (which fosters idea exchange) and collective intelligence. Audretsch and Feldman (1996) showed that industrial innovation is more likely to occur in areas with relatively large knowledge externalities, meaning the activities of agents in those areas are likely to enhance the knowledge of others with close technical or geographic proximities. Jaffe (1986) showed that the creative productivity and the research and development returns of firms correlated positively with their involvement in research-intensive technology groups, groups in which they could exchange ideas. Orlando and Verba (2004) illustrated that industrial innovation is more common in urban areas, areas that foster significant idea exchange. Perry-Smith and Shalley's (2003) research showed that social interactions influence individual's and organization's abilities to innovate, thus also supports the notion that meme exchange influences collective intelligence. This position is also supported by Tsai's (2001) analysis of 24 petrochemical companies and 36 food manufacturing companies showing that organizational units are more effective at stimulating innovation when they are centrally located in the company and, in turn, foster increased idea exchange.

In order to maximize their research potential educational institutions and other organizations must recognize that groups of people are more intelligent than any individual. Not only is this important, but it grows increasingly so with time. The birth of the internet led to a rapid growth of collective intelligence, a growth explained by Metcalfe's Law, a law that states that the value of a network is proportional to the square of the number of processors and users in it (Gilder, 2000). Kurzweil (2006) took this idea even further suggesting that human culture is approaching the knee of the exponential curve of its growth, the point at which it becomes apparent that our rate of intellectual and technological developments are for all practical purposes, approaching infinity.

2.7 Summary

The new paradigm of complexity science offers a plethora of theoretical and methodological tools that researchers can use to analyze the intricate mechanics of social systems. These tools and theories offer new possibilities for researchers to solve problems by circumventing the limitations of traditional linear analysis techniques. This paradigm illuminates the processes by which complex social systems emerge out of networks of interacting individuals. The evolution of ideas, memes (Dawkins, 1976), is one of these emergent phenomena.

Ideas move from person to person as people communicate and interact with each other. The ideas that enable modern society to operate are complex systems of multiple co-dependent memes, memeplexes. Together these memes form solutions to real world problems and, in turn, incentivize people to adopt them. By forming memeplexes that solve problems making people's lives better, the selfish memes ensure their mutual evolutionary survival.

In academic systems and other collaborative settings, good ideas and new problem solutions spread as researchers adopt ideas from their colleagues. Sometimes a researcher will adopt an idea, make a small alteration to it, and pass it on to a collaborator. The memeplex was adopted, some of the component memes were altered, and the memeplex was passed on to someone else. The memeplex evolved through the processes of asexual reproduction and mutation. These direct lineages of meme replication and mutation trace back to the earliest humans. However, there is also another evolutionary mechanism at work. On occasion an individual will take parts from two problem solutions, memeplexes, and recombine them to form a new and unique memeplex. This recombination of memes parallels the recombination of genes in sexual reproduction and accelerates the rate of creative evolution.

The sexual reproduction of memes provides a new explanation for collective intelligence, the phenomenon where groups of people are more intelligent than any of their constituent individuals. By modeling the asexual and sexual reproduction of memeplexes in collaborative networks, this study intends to examine the conditions under which problem solutions evolve. By examining this evolutionary process the model should provide insight into the conditions that optimize the capacity of various organizations and institutions to solve complex, real world problems.

CHAPTER 3

RESEARCH DESIGN

In order to help explain the phenomenon of creative evolution, this computational model was programmed in GNU Octave (Eaton, 2013). GNU Octave is a high-end programming platform that is an open source version of MATLAB. The platform has built in functions that enable programmers to easily implement complex mathematical calculations and algorithms into their code and enables users to easily manipulate and produce graphical data outputs (Mathworks, 2013). This model runs an agent based simulation of memetic evolution. Bonabeau (2002) stated, "in an agent based model a system is modeled as a collection of autonomous decision-making entities called agents. Each agent individually assesses its situation and makes decisions on the basis of a set of rules" (p. 7280). The concept of agents that interact to produce complex behaviors existed well before the modern computer and can be traced back as far as Adam Smith's (1776) description of individuals as selfish entities, the agents from which he deduced his economic conclusions. In this model the agents are computer representations of the cognitive "selves" described in Chapter 2. As consistent with the philosophic view of objective reality outlined in the first two chapters the model was designed to represent a specific predefined problem that the agents at hand are working to solve. Miller and Page (2007) stated,

> One of the most powerful tools arising from complex systems research is a set of computational techniques that allow a much wider range of models to be explored. With these tools, any number of heterogeneous agents can interact in a dynamic

> environment within the limits of time and space. Having the ability to investigate new theoretical worlds obviously does not imply any kind of scientific necessity or validity—these must be earned by carefully considering the ability of the new models to help us understand and predict the questions that we hold most dear (p. 5).

The remainder of this chapter is dedicated to establishing the validity towards which, Page and Miller (2007) explain we must carefully strive. This validity is established by summarizing the computational techniques used. The most important step in establishing the validity of the code was the establishment of an array of design requirements that the model had to meet. Here, the study contained an outlined list of philosophical, sociological, and practical assumptions made as a result of the literature review and arguments constructed in Chapter 2. These assumptions were then translated into the design requirements that served as constraints to guide the model's development.

3.1 Predictive and Explanatory Models

Modern computational models come in two varieties: predictive and explanatory, however, the difference between the two is actually quite subtle. Predictive models use observations of past phenomena to develop algorithms and equations that will predict the behaviors of similar phenomena in the future. On the other hand, explanatory models do not seek to predict any particular future phenomena. Rather, they attempt to explain the logical dynamics of a larger set of related phenomena. For example, if we discuss the relationship between parental incomes and student SAT scores, it might be possible to develop a computational or mathematical model to predict future student SAT scores based on parental incomes and a variety of other input variables (with a degree of variance). Here, it is important to note that this example models correlations in the relationship between SAT scores and a variety of other variables without providing any insight into the causal relationships between the variables. By sacrificing the predictive abilities of models for more generalizable explanatory abilities we can gain better insights into various causal relationships. An explanatory model would not provide us with the ability to predict SAT scores, however its causal insights would likely have more power to help us think about other exams, such as the ACT.

After developing an explanatory model the researcher should extend their analytical lens beyond the model and examine it within a framework of related theory. By doing this the researcher can develop an explanatory model with less specificity but more generalizability than its predictive counterpart (Page & Miller, 2007). Once this step is finished the researcher would then have a logical justification for deciding upon which future empirical work to complete. Due to the abstract theoretical nature of this study's simulations in addition to its broad range of assumptions it is unreasonable to assume this model can predict the outcome of complex creative systems, therefore it is only intended to be explanatory. This approach is consistent with Page and Miller's (2007) recommendation to focus on the general theory rather than the specific model mechanics.

3.2 Machine Learning: Navigating Fitness Landscapes

Machine learning is a relatively modern computational technique that is used to find patterns in data (Mohri, Rostamizadeh, Talwalkar, & Ebrary, 2012**)**. Common examples include email spam filters (Mohri et el., 2012), which use databases of old spam emails to search for patterns which will help filter spam from your inbox with the highest level of efficiency possible (Nelson, Barreno, Jack, Anthony, Joseph, Rubinstein, Saini, Sutton, Tygar, & Xia, 2008). The following study employs two types of machine learning algorithms, hill climb algorithms and genetic algorithms.

Hill climb algorithms represent a subset of machine learning techniques in which a computer program works to evolve a single solution to a provided problem (Mitchell, 1997). This evolution generally takes place with respect to some mathematical function that has been designed to test the quality of the solution as a function of the solution parameters (Holland, 1974; Mitchell, 1997). These output functions can be mathematically derived from mathematical or physical principles. An example of this might occur when trying to optimize the torsional load a particular beam can handle given certain geometric constraints.

In other examples these algorithms may be used to optimize the value of a more complex function that applied weighted values to various different attributes of a problem solution. The program enters either a random starting value for each input variable or a predetermined starting value for each input variable. The model in this study assigns random starting values, as it assumes no prior knowledge of the problem at hand. These values are assessed for quality as determined by the program's quality assessment function. Once this process is completed, a pseudo-random number generator is used to make pseudo-random changes to a variety of the solution's input values. Now this new mutated design is run through the same quality assessment function and a new quality value is calculated. Here, the algorithm utilizes an if-then statement that presents two possible courses of action. If the new mutated design is assessed as being higher quality than the previous iteration the mutated design will go through a second round of pseudo-randomly generated mutations and the process will continue. On the other hand, if the new mutated design is assessed as being lower quality than the previous iteration, then the design reverts to the previous iteration's settings and goes through a new set of pseudo-randomly generated mutations. The process continues. This process can either be set to run through a set number of iterations or until the output quality value meets some predefined criterion.

Machine learning techniques can be eloquently illustrated through a series of geometric examples. Suppose that the number of input variables for a given machine learning system is represented by the variable n. Now, let us picture a graphical representation of this system of $n+1$ dimensions where the first n dimensions represent the input variables and the final dimension represents the output variable. A simple three-dimensional example of this can be visualized with traditional Euclidean thinking. Consider an example with two input variables. Here we can imagine a contour plain with a very similar appearance to that of a map with proper dips and bumps to show elevation. The x and y axes, which run parallel to the lines of latitude and longitude, respectively, correspond with the values of the two input variables. Now one can consider a particular combination of x and y coordinates and imagine the corresponding spot on the surface. Naturally, this point will also have an elevation, or z-coordinate which represents the output variable. While it is not always true that a larger output value, a larger z-coordinate, is higher quality than smaller z values, for the sake of clarity we will assume that in this case it is true that larger values are better (in this study's model it is the opposite). Here, we can think of the hill climb algorithm as doing exactly what the name suggests. The algorithm mutates the solution's x and y values, keeping them when they improve (or increase) the elevation, reverting when they do not, and iterating upon this process until the solution climbs to the highest peak, the point illustrating the optimal design quality value.

Now, imagine an example where the quality (output) value shows a shape created by a concave down parabola rotated about the z-axis. Here, the parabolic hill-shaped surface has one peak, representing a single optimum, thus, a hill climb algorithm will easily be able to navigate to the optimal solution, the peak. Genetic algorithms, a more complex form of machine learning algorithm, are unsuited for this situation and better utilized on fitness landscapes with multiple peaks. Ultimately, utilizing a genetic algorithm in this case would require the computer to run unneeded operations and slow down the program's optimization process. Hill climb algorithms are apt to solve problems with a limited number of peaks or optima (Holland, 1974; Mitchell, 1997).

For purposes of comparison, let us consider an example with 26 input variables. While the human brain cannot picture a 27 dimensional surface, this non-Euclidean surface is still representative of the idea at hand. The number of possible directions a hill climber could move in his search for a peak can be calculated using Equation 1. D is the number of directional possibilities, and n is the number of input variables.

$$D = 2^n \tag{1}$$

Using this equation we see that the two input example yields four directional combinations in which the hill climber can try moving. Contrarily, the 26 input example yields 67,108,864 possible directional combinations. These surfaces are examples of fitness landscapes, multidimensional surfaces with a dimension for each input and one for the output. Agents wander the surfaces in order to find the set of inputs that provide the optimal output. When using machine learning algorithms to search surfaces, agents' abilities to successfully find the ideal output are path dependent, meaning, their future success is dependent upon their previous decisions (Page, 2006). An agent's choice to move in the wrong direction along a fitness landscape during a past iteration may prevent it from ever finding the global optimum, the point on the fitness landscape with the absolute best output value. Rather, the agent may end up stuck at a local optimum, a point with a high output value relative to local areas but not on the global (entire) fitness landscape.

Just like hill climb algorithms, genetic algorithms (Holland, 1975; Mitchell, 1997) are designed to optimize problem parameters with respect to some designated problem quality function. The primary difference between hill climb algorithms and genetic algorithms is that genetic algorithms iterate upon a pool of potential design solutions, rather than a single design solution lineage, as is the case with hill climb algorithms. In genetic algorithms the multiple design solutions run through the iterative mutation quality assessment cycle that is used in hill climb algorithms; however, genetic algorithms implement a second mechanism to the evolutionary process. After each problem solution in the idea pool is pseudo-randomly mutated the design inputs go through a crossover process (Holland, 1974; Mitchell, 1997). In this crossover process, a few pseudo-randomly selected inputs in each problem solution are switched with inputs from another pseudo-randomly selected problem solution.

Another difference between hill climb algorithms and genetic algorithms lies in the selection mechanism that determines whether a particular design solution will survive until the next iteration. Instead of simply testing the fitness of a single solution and comparing it to the previous iteration's fitness, genetic algorithms cull a certain portion of the population's solutions with the weakest fitness values, allowing only the strong to survive. Some of the surviving solutions are cloned in order to keep the design population at a constant size. While hill climb algorithms follow the evolutionary patterns of asexually reproducing genes and memes, genetic algorithms follow the evolutionary patterns of sexually reproducing genes and memes. Mitchell (1997) showed that alternating the use of hill climb and genetic algorithms can be used to optimize the effectiveness of machine learning optimization processes.

Other artificial intelligence mechanisms have been used to simulate and model social and cognitive systems. One of these systems, neural networks, was actually designed to replicate many of the feedback systems in the brain; however, machine learning techniques were used in this study for two primary reasons. First, machine learning is not a black box process, but artificial neural networks are (Haykin, 1994). Black box processes only allow users to see the process inputs and outputs and prevent users from observing intermediate stages in their behavior. Miller and Page (2007) recommended against the use of black boxes. Second, machine learning algorithms were chosen because they parallel the iterative nature of the research process.

3.3 Design Requirements

Design requirements illustrate a variety of conditions that the computational model must meet in order to validate that it, in fact, models the phenomena that it is designed to model. These design requirements are broken into two parts. First, an assumption is provided which illustrates a specific statement about the nature of creative systems. In parentheses at the end of each of these assumptions is the number of the section that reflects and discusses the justification listed. Immediately following each of these assumptions the actual design requirement is listed, showing exactly what the computational model must do in order to validate that it meets its respective assumption.

3.3.1 Complex Reality

Assumption: Real world problems are complex in nature. Solutions are dependent upon the complex interactions between different problem components. (Section 1.6)

Design Requirement: The designed problem must be non-linear.

3.3.2 Continuous Solution Quality

Assumption: Solutions to real world problems are not simply right or wrong; they are better or worse. (Section 1.6)

Design Requirement: Problem solutions must not provide binary results.

3.3.3 Agent Preferences

Assumption: Individuals all seek the same ends when working in specialized areas. This means that all individuals interpret results in the same manner and that problems have definable goals, such as obtaining equal SAT scores between different socioeconomic, racial, and ethnic groups or reducing the number of automobile fatalities in a given car model to zero. (Section 1.4)

Design Requirement: All agents receive the same feedback values when they input the same values into the simulated problem.

3.3.4 Objective Problem Goals

Assumption: Humans seek the best possible solution to the problem on which they are working. For example, we assume that cancer researchers unanimously want to destroy cancer. (Sections 1.5)

Design Requirement: Agents will never revert to a past problem solution once they have generated a new solution which yields a superior feedback.

3.3.5 Asexual Creativity

Assumption: Humans are able to creatively seek new solutions to problems, without communicating with other agents. They do this by randomly making changes to past solutions, through mechanisms that parallel evolution in an asexually reproducing species (Sections 2.1 & 2.3).

Design Requirement: Agents must be able to randomly change portions of problem solutions in order to model asexual modes of creativity.

3.3.6 Sexual Creativity

Assumption: Humans are also able to creatively seek solutions to problems by communicating and sharing ideas with other humans, through mechanisms that parallel evolution in a sexually reproducing species. (Sections 2.1 & 2.4)

Design Requirement: Agents must be able to copy portions of problem solutions from other agents in order to model sexual modes of creativity.

3.3.7 Unique Problems

Assumption: Humans are capable of using their creative capacities to pursue solutions to a wide variety of complex problems. (Section 1.4)

Design Requirement: All mathematical problem coefficients must be able to be randomly re-generated.

3.3.8 Boldness of Asexual Creativity

Assumption: Humans can alter different percentages of old problem solutions when attempting to asexually create new and improved problem solutions. (Section 2.3)

Design Requirement: The model must be able to adjust the percentage of an old idea that an agent changes when changing it in attempt to find a better problem solution. (Research sub-question 3)

3.3.9 Boldness of Sexual Creativity

Assumption: Humans can copy different percentages of other people's ideas when attempting to sexually create new and improved problem solutions. (Section 2.4)

Design Requirement: The model must be able to adjust the percentage of an idea one agent copies from another in attempt to find a better problem solution. (Research sub-question 4)

3.3.10 Creative Time Allocation

Assumption: Humans may spend different amounts of time pursuing creative solutions to problems individually as opposed to collaboratively (Section 2.4)

Design Requirement: The model must be able to adjust the frequency with which agents utilize asexual modes of creativity as opposed to sexual methods of creativity. (Research sub-question 1)

3.3.11 Leaders

Assumption: Humans can choose which agents to copy solutions from based upon different criteria. (Section 2.4)

Design Requirement: The model should allow agents to either copy solution components from the agent who currently holds the best problem solution output or from an agent selected at random. (Research sub-question 2)

3.4 Design Guidelines

The following design guidelines are different from design requirements in that they represent desirable characteristics of the model, rather than functions that the model must have in order for it to maintain validity. These guidelines include hardware guidelines based upon the desire to create a technically feasible and efficient computational model. These guidelines are also used to assure that the model's function is clearly communicated so that future researchers can easily edit it as necessary. In addition, these requirements are used to assure the simulation consistently outputs data in a clear and concise form.

3.4.1 Researcher Usability

This program should be designed to run on as small of a processor and as little RAM as possible in order to maximize the availability of this resource to researchers with varying computational limits; however, the ability of the model to accurately reflect the phenomena at hand should take precedence over this guideline.

3.4.2 Code Clarity

The code must be properly commented and include instructions describing how users may modify the code in order to experiment with the model (McConnell, 1993).

3.4.3 Output Clarity

This program should produce the clearest outputs possible (McConnell, 1993).

3.4.4 Code Simplicity

All else being equal, simpler models are preferable (Miller & Page, 2007).

3.5 Breaking Down the Code

This section illustrates various features of the computational model in order to show how the program meets all of the design requirements. The following table illustrates which of the subsections provides the justifications for each of the design requirements.

Table 1

Design Requirement Validation Table

Design Requirement	Design Requirement Name	Validation Section
3.3.1	Complex Reality	3.5.1
3.3.2	Continuous Solution Quality	3.5.1
3.3.3	Agent Preferences	3.5.1
3.3.4	Objective Problem Goals	3.5.1
3.3.5	Asexual Creativity	3.5.2
3.3.6	Sexual Creativity	3.5.2
3.3.7	Unique Problems	3.5.3
3.3.8	Boldness of Asexual Creativity	3.5.4
3.3.9	Boldness of Sexual Creativity	3.5.4
3.3.10	Creative Preferences	3.5.4
3.3.11	Leaders	3.5.4

Note: This table shows where the validations for each of the design requirements can be found.

3.5.1 Modeling Complex Problems (Design Requirements 3.3.1-3.3.4)

The externally defined problem in the code consists of an idea represented by 25 digits between zero and nine. This idea could be viewed as a representation of an ancient human hand axe design, a modern automobile headlight design, or even a larger conception of some societal ideal as defined to a specific measurable metric. Each of the 25 digits can be viewed as one of 25 subcomponent portions of the larger design, one of the 25 memes in the memeplex. These 25 digit ideas are randomly determined by the code and represent the ideal values for all aspects of the problem solution.

Two input variables are pseudo-randomly generated at the beginning of the simulation. Each agent starts with a 25-by-1 matrix, a row of 25 randomly generated integers between zero and nine identified by the variable *A*. These are the initial conditions of the 25 solution conceptions to 25 problem subcomponents. Likewise, a single 25-by-1 matrix of randomly generated integers between one and nine is created representing the ideal solution to the simulated problem and is identified by the variable *B*.

Solution quality is not determined linearly, meaning it is not determined by the influences of the 25 solution subcomponents independently. Solution quality, instead, is determined by the interactions between those subcomponents. For example the quality of a soda's taste (to the average consumer) is not simply dependent upon the sodium, sucrose, and flavor contents independently, but dependent upon the interactions between those components. An increase in the concentration of one flavor might only be a good thing if the concentrations of other flavors are decreased. The interactions between these solution subcomponents are calculated in a new 25-by-25 matrix, *G*, by finding the products of every possible pair of problem solution subcomponents as shown in Equation 2.

$$G = A'A \tag{2}$$

Likewise, a 25-by-25 matrix, *I*, is calculated giving the products of every possible pair of ideal solution subcomponent interactions as shown in Equation 3.

$$I = B'B \tag{3}$$

Redundancies are addressed in the square matrices, *G* and *I*. These redundancies are illustrated by Equation 5 and Equation 6 where the index variable x indicates the column and index variable y indicates the row.

$$G_{xy} = G_{yx} \tag{4}$$

$$I_{xy} = I_{yx} \tag{5}$$

In order to eliminate these redundancies, upper triangularizations are taken for both matrices *G* and *I*, creating two new matrices, G_U and I_U. This process replaces all values below the diagonal where

$$x = y \tag{6}$$

with zeros, removing all redundant values. Here, the values along the diagonal where the row number is equal to the column number are not set to zero. This diagonal represents the product of each possible problem solution or ideal solution subcomponent with itself, its square. This diagonal is left in place illustrating that while problem solution quality is not exclusively dependent upon the solution subcomponents independently, their independent values might have some impact on the overall solution quality.

Interaction values, the products of every possible pair of solution subcomponents for both the agent's solution and the ideal solution, are represented in the matrices G_U and I_U respectively. The solution proximity, or the similarity between the agent's solution and the ideal solution for all subcomponent interaction values, is calculated in the new matrix, *P*, as illustrated in Equation 7.

$$P = I_U - G_U \tag{7}$$

Not all aspects of a problem solution have equal influence on the ability of a solution to solve the problem at hand. A 25-by-25 matrix, W, was generated at the onset of the simulation and randomly filled with weighting coefficients, randomly assigned integers between zero and nine. Equation 8 calculates the weighted interaction values, P_W, by taking the Hadamard product of matrices W and P, creating a 25-by-25 matrix.

$$P_W = W \circ P \tag{8}$$

Solution quality is not linearly dependent upon the weighted, ideological proximity of different subcomponent interactions. In the field of statistics, error rates are often found using the sum of squared differences, not the sum of differences (Draper & Smith, 1998). A specific subcomponent interaction being twice as far from the ideal solution value would likely reduce the quality of the solution at hand by more than a factor of two. Imagine a person designing a spear. Here, it makes sense that a spear, which is slightly dull, might be somewhat effective despite it not being perfectly ideal, but as the dullness of the spear approaches some limit (the edge moves away from the ideal sharpness level) the functionality of the solution diminishes increasingly. This model feature is implemented by squaring all of the weighted solution subcomponent interactions inside matrix, P_W, to develop a new matrix F as seen in Equation 9 which is found by taking the Hadamard product of matrix P_W with itself.

$$F = P_W \circ P_W \tag{9}$$

At this point F contains values that are directly proportional to the effect each design subcomponent interaction has on the overall solution quality. The aggregate of all of these component effects determines the overall solution quality. Here, the overall solution quality is determined by calculating the summation of all of the design quality components, as illustrated in Equation 10.

$$Q = \sum_{x=1}^{25}\left(\sum_{y=1}^{25} F_{xy}\right) \tag{10}$$

The value Q produced by Equation 10 represents the final feedback that an agent receives upon testing a problem solution against the theoretical ideal solution. This value decreases as the various subcomponents of the agent's solution approach their corresponding ideal solution values. The agent interprets this change as positive and uses the value of Q as the fitness test criterion in every iteration of this simulation, constantly working to improve its problem solution in order to further decrease the returned feedback value, Q.

3.5.2 Modeling Creativity with Machine Learning (Design Requirements 3.3.5-3.3.6)

In order to simulate creative evolution an agent simply repeats this inductive, trial and error sequence for the desired number of iterations. After each iteration, the basic fuel of evolution, mutation is added. Random changes are made to some of the randomly selected 25 idea conception components. Having made these random changes, the agent then tries its new idea solution. If this new idea produces a lower (more ideal) solution feedback value, then the agent keeps its altered 25 digits. During the next iteration it mutates again and repeats the process. If the agent does not yield an enhanced solution value, then it will revert to the previous solution and repeat the process during the next iteration.

At this point the section has explained how the inductive model simulates asexual modes of creativity but has not explained the sexual mechanisms. Here, a second process was implemented into the model in order to parallel the sexual analog of the asexual mode of creativity. In order to do this the code was designed so that it operates with 25 agents attempting to solve the problem rather than one. In order to fix the problem of isolation between agents a function was added into the program that gave agents an alternative creative method to the traditional asexual process described above. The agents were programed to copy some of a randomly selected agents', randomly selected 25 idea components. Once again after taking on their new forms, these agents would try out their new problem solutions, keep the new ideas if they improved their solution feedback values, and revert if not.

3.5.3 Modeling Unique Problems (Design Requirement 3.3.7)

Pseudo-random number generators are found throughout the body of this code. A pseudo-random number generator is a type of algorithm that enables computers to simulate the generation of random numbers within a defined range (Gentle, 2004). These generators have two functions in this model. First, they are used to generate the random mutations of components of random agents' 25 digit ideas. Secondly, they are used to generate the coefficients used to represent particular aspects of particular problem solutions, the 25 ideal value digits, as well as each of the agents' initial 25 guess digits.

This occurs at the onset of the simulation and is essential to assure the findings of this research represent creative evolution as a whole rather than creative evolution with respect to one randomly generated problem.

3.5.4 Optimizing the Creative Cycle (Design Requirements 3.3.8-3.3.11)

Miller and Page (2007) recommended implementing “tunable dials” (p. 248) into computational models that can be easily manipulated to alter the parameters of the simulations. For this reason the simulation was designed to enable easy constraint alteration to allow the simulation independent variables to be easily changed and to make the model easily adaptable in case of future research. Four independent variables were changed throughout the simulations for this study, paralleling the four research sub-questions. First, the agents were either set to copy the memes of the best agent with the best solution in the group or meme of an agent selected at random. This variable is referred to as the leader (L.), with a best and random setting.

Next, the extent to which the agents manipulated their memeplexes, their 25 solution digits, when attempting to solve problems in isolation had two possible settings. For the low setting, a randomly generated number between one and seven of the agents’ 25 meme digits were mutated. For the high setting, a randomly generated number between one and eighteen of the agents’ 25 meme digits were mutated. This variable is referred to as the asexual level (A.L.). Likewise, the number of memes agents copied from other agents’ memeplexes had two possible settings. For the low setting agents copied a randomly generated number between one and seven random memes from another agent’s memeplex. For the high setting agents copied a randomly generated number between one and eighteen random memes from another agent's memeplex. This variable is referred to as the sexual level (S.L.). The numbers seven and eighteen are arbitrary values for the low and high settings of these input variables, yet this is justifiable as this study is a first trial and smaller increments can be selected in the future so as to examine this data in greater resolution.

The only continuous independent variable manipulated throughout the simulation trial was the probability of an agent employing sexual modes of creativity as opposed to the asexual counterpart. These values were set between 0 percent and 100 percent at 20 percent intervals. This variable is referred to as the sexual percent (S.P.). For the sake of clarity and organization the four independent variables, an explanation of each, and their abbreviations are listed in Table 4. In total, this yields 48 different combinations of input variables. For each of these 48 combinations of input variables, two hundred simulations were run, and each simulation was run for one hundred creative iterations. If any agent discovered the perfect problem solution within these one hundred iterations then the simulation was deemed a success. Otherwise, it was deemed a failure.

Table 2

Independent Variables

Variable	Explanation	Abbreviation
Sexual Percent	How often do agents exchange ideas?	S.P.
Sexual Level	How much of another agent's idea does an agent copy?	S.L.
Asexual Level	How much of its idea does an agent mutate?	A.L.
Leader	Whose ideas do agents copy?	L.

Note: This table lists the independent variables manipulated in this study.

3.6 Data Analysis

The outcomes of the 9,600 simulations in the study were analyzed using a method know as Chi-squared Automatic Interaction Detector (CHAID), a tree based statistical method proposed by Kass (1980). CHAID builds decision trees with multiple branches based on an algorithm that splits the branches according to statistically significant differences in success rates. Here, the tree can contain a number of splits equal to the number of independent variables. The order in which these splits branch out from the base of the tree occurs in the order that the independent variables correlate with the largest differences in the dependent variable, from greatest to least. This method is commonly used when splitting data sets into many multi-way frequency tables where the success ratio of a categorical dependent variable is affected by multiple independent variables. CHAID utilizes the Bonferroni correction method in order to reduce the occurrence of type I errors as a result of multiplicity. Multiplicity causes type I errors because the large number of independent variables leads to a large numbers of null hypotheses (Statsoft Inc., 2013).

CHAPTER 4

FINDINGS

A total of 9,600 simulations were run in this study. The simulations accounted for all 48 combinations of the four independent variables. As discussed in Chapter 3, each trial simulated 25 agents working together to solve a complex problem for 100 creative iterations. These problems defined the ideal solution value as 0, thus lower values were deemed better.

4.1 A Sample Simulation

Figure 1 illustrates a graphical representation of a sample simulation. In this particular example the leader was set to best, the asexual level was set to high, the sexual level was set to low, and the simulation agents used sexual modes of creativity 0 percent of the time. The graph illustrates the best solution value yielded by any of the 25 agents for the 100 creative iterations the simulation ran. Due to the large changes in the returned quality function (multiple orders of magnitude) during the simulation, the solution quality is illustrated on a logarithmically scaled y-axis. In this particular simulation the solution quality was still above 1,000 when it terminated after creative iteration 100, showing that in this particular trial no agent reached the optimal value of 0.

Another phenomenon of note is illustrated in the graph by the repeating pattern of horizontal lines dissected by spots where the values quickly decrease and settle at new horizontals. These horizontal portions of the graph illustrate portions of the simulation where the agents creatively stagnated. In these flat line segments, 0 improvements were made to the best solution value (the value saw no decreases) so the line remained horizontal. These sections of creative stagnation are the tentative optima of the simulation or the best problem solutions the collaborative system had found up to that point but not the truly ideal solutions. The sharp decreases illustrate points where one of the simulation's 25 agents finally managed to yield a lower (better) problem solution. While Figure 1 illustrates that during significant portions of the 100 iteration creative simulation the agents did not improve their best solution, this does not mean that no agents yielded improvements to their respective solutions.

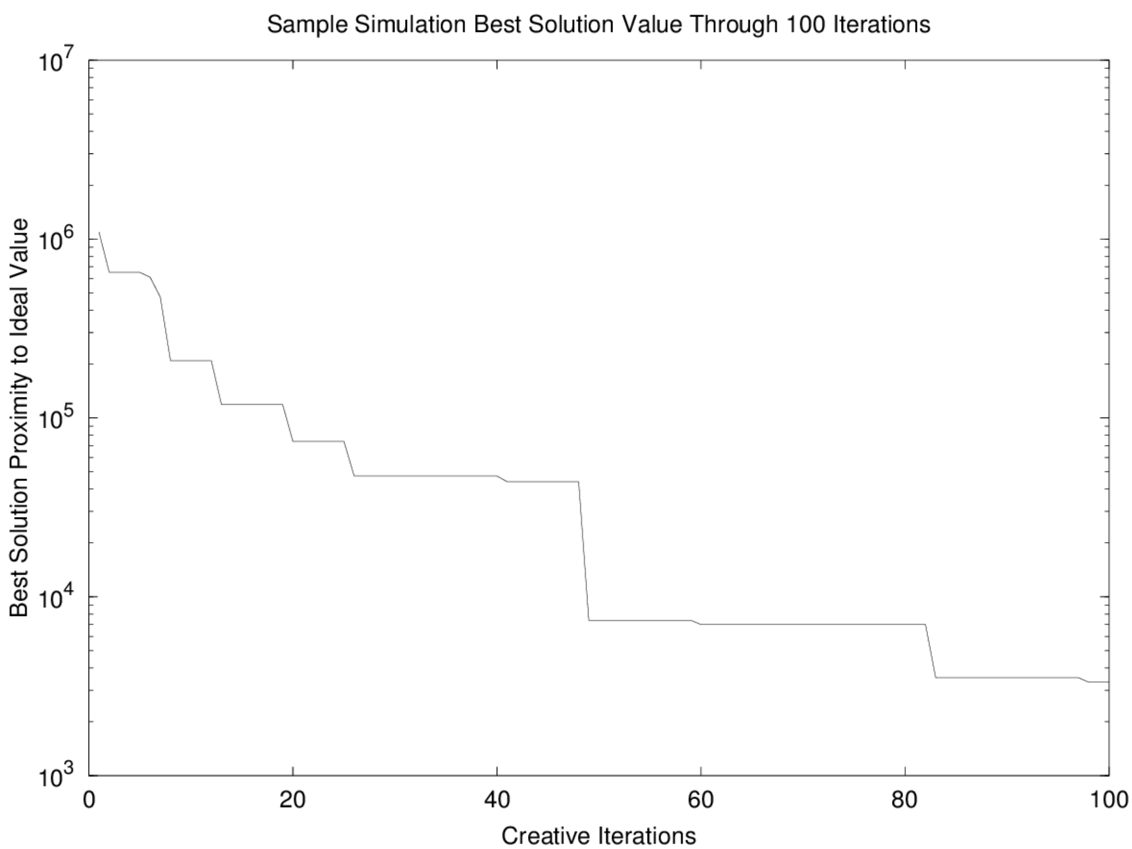

Figure 1. Sample simulation graph of best agent solution feedback value through 100 creative iterations.

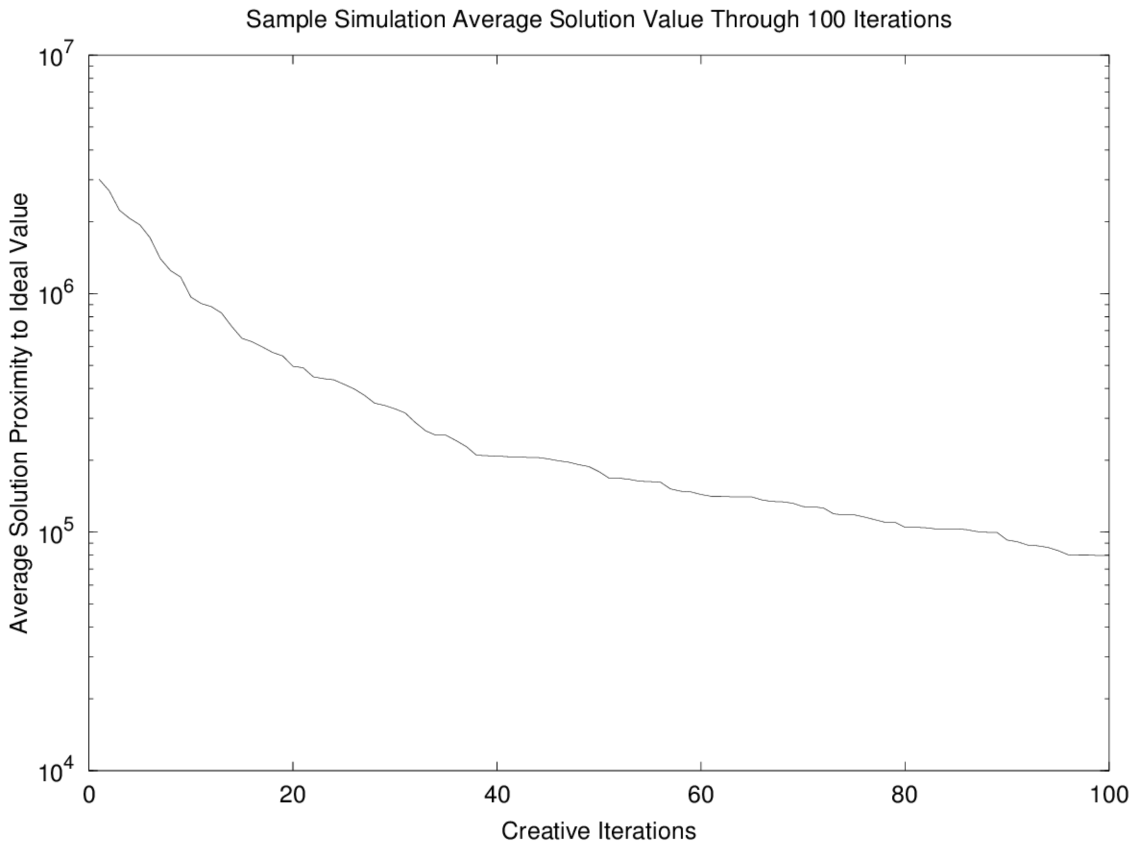

Figure 2. Sample simulation graph of average solution value through 100 iterations.

The lack of horizontal sections in Figure 2, which illustrates the average of all 25 agent solution values, shows how the agents can yield improvements in their own solutions even when the best solution value is not improving.

4.2 Results

A summary of the 9,600 simulation outcomes can be seen in Table 3. The table shows the number of successes and failures out of the 200 trials run for each of the 48 different combinations of input variables, successes being defined as simulations in which one of the 25 collaborating agents reached the ideal problem solution within 100 creative iterations.

Table 3

Simulation Results

Leader	Asexual Level	Sexual Level	Sexual Percent	Successes	Failures
Random	Low	Low	0%	51	149
Random	Low	Low	20%	155	45
Random	Low	Low	40%	199	1
Random	Low	Low	60%	195	5
Random	Low	Low	80%	188	12
Random	Low	Low	100%	136	64
Random	High	Low	0%	24	176
Random	High	Low	20%	122	78
Random	High	Low	40%	188	12
Random	High	Low	60%	186	14
Random	High	Low	80%	178	22
Random	High	Low	100%	120	80
Random	Low	High	0%	49	151
Random	Low	High	20%	181	19
Random	Low	High	40%	200	0
Random	Low	High	60%	193	7
Random	Low	High	80%	179	21
Random	Low	High	100%	102	98
Random	High	High	0%	16	184
Random	High	High	20%	172	28
Random	High	High	40%	190	10
Random	High	High	60%	181	19
Random	High	High	80%	166	34
Random	High	High	100%	94	106
Best	Low	Low	0%	56	144
Best	Low	Low	20%	178	22
Best	Low	Low	40%	194	6
Best	Low	Low	60%	188	12
Best	Low	Low	80%	162	38
Best	Low	Low	100%	45	155
Best	High	Low	0%	24	176
Best	High	Low	20%	170	30
Best	High	Low	40%	177	23
Best	High	Low	60%	160	40
Best	High	Low	80%	128	72
Best	High	Low	100%	53	147
Best	Low	High	0%	59	141
Best	Low	High	20%	190	10
Best	Low	High	40%	194	6
Best	Low	High	60%	180	20
Best	Low	High	80%	136	64
Best	Low	High	100%	8	192
Best	High	High	0%	23	177
Best	High	High	20%	170	30
Best	High	High	40%	164	36
Best	High	High	60%	135	65
Best	High	High	80%	95	105
Best	High	High	100%	8	192

Note: The table shows the number of successes and failures for 100 trials for each of the 48 combinations of inputs.

Table 4

CHAID Analysis

All Simulations	Branch Level 1	Branch Level 2	Branch Level 3	Branch Level 4
66.3%; Next Split Criterion: S.P. (*P*=0.000)	S. P.=0%; **81.1**%; N. S. Criterion: A.L. (*P*=0.000)	A.L.=Low; **26.9%**;		
		A.L.=High; **10.9%**;		
	S.P.=20%; **83.6%**; N. S. Criterion: S. L. (*P*=0.000)	S.L.=Low; **78.1%**; N. S. Criterion: L. (*P*=0.000)	L=Random; **59.2%**; N. S. Criterion: A.L. (*P*=0.000)	A.L.=Low; **77.5%**
				A.L.=High; **61.0%**
			L.=Best; **87.0%**	
		S.L.=High **89.1%**; N. S. Criterion: A.L. (*P*=0.001)	A.L.=Low; **92.8%**	
			A.L.=High; **85.5%**	
	S.P.=40%; **94.1**%; N. S. Criterion: A. L (*P*=0.000)	A.L.=Low **98.4%**; N. S. Criterion: L. (*P*=0.002)	L.=Random; **99.8%**	
			L.=Best; **97.0%**	
		89.9%; N. S. Criterion: L. (*P*=0.000)	L.=Random; **94.5%**	
			L.=Best; **85.2%**	
	S.P.=60%; **88.6%**; N. S. Criterion: A.L. (*P*=0.000)	A L. = Low **94.5%**; N. S. Criterion: L. (*P*=0.002)	L.=Random; **97.0%**	
			L.=Best; **92.0%**	
		A.L.= High **82.8%**; N. S. Criterion: L. (*P*=0.000)	L.=Random; **91.8%**	
			L.=Best; **73.8%**; N. S. Criterion: S.L (*P*=0.004)	S.L.=Low; **80.0%**
				S.L.=High; **67.5%**
	S.P.=80%; **77.0%**; N. S. Criterion: L. (*P*=0.000)	L.=Random; **88.9%**; N. S. Criterion: A.L. (*P*=0.010)	A.L.=Low; **91.8%**	
			A.L.=High; **86.0%**	
		L.=Best; **65.1%**; N. S. Criterion: A.L. (*P*=0.000)	A.L.=Low; **74.5%**; N. S. Criterion: S.L. (*P*=0.003)	S.L.=Low; **81.0%**
				S.L.=High; **68.0%**
			A.L.=High; **55.8%**; N. S. Criterion: S.L. (*P*=0.001)	S.L.=Low; **64.0%**
				S.L.=High; **47.5%**
	S.P.=100%; **35.4%**; N. S. Criterion: L. (*P*=0.000)	L.=Random; **56.5%**; N. S. Criterion: S.L. (*P*=0.000)	S.L.=Low; **64.0%**	
			S.L.=High; **49.0%**	
		L.=Best; **14.2%**; N. S. Criterion: S.L. (*P*=0.000)	S.L.=Low; **24.5%**	
			S.L.=High; **4.0%**	

Note: This table shows the decision tree produced by CHAID. Various abbreviations are used for independent variables including Sexual Percent (S.P.), Asexual Level (A.L.), Sexual Level (S.L.), and Leader (L.). The bold boxes correspond to the decisions yielding optimal success rates.

The simulation results were then analyzed using CHAID in SPSS (IBM Corp., 2013) generating the decision tree illustrated in Table 4. This table shows which independent variables caused significant changes in the simulation success ratios. All branches in the tree were found to be statistically significant with the lowest significance split having an alpha value ($p = 0.010$), five times lower than the traditional test statistic ($\alpha = 0.050$). The one continuous variable, the percentage of the time in which the simulation agents engaged in sexual idea exchange as opposed to asexual idea generation, played the most significant role in influencing the problem solving success ratio. Here, the data were split into six percentage ranges. Simulation success ratios for values between 40 percent and 60 percent were the highest. Independent values further away from this middle range, 40 percent to 60 percent, on both the high and low ends yielded worse simulation success ratios. These values are illustrated in Table 5 and graphically represented in Figure 3.

Table 5

Success Occurrences for Different Idea Exchange Frequencies

Idea Exchange Freq.	Successes	Failures	Success Percent
0%	302	1298	18.9%
20%	1338	262	83.6%
40%	1506	94	94.1%
60%	1418	182	88.6%
80%	1232	368	77.0%
100%	568	1034	35.4%

Note: The table shows the number of successes, failures, and success percentages for each idea exchange frequency

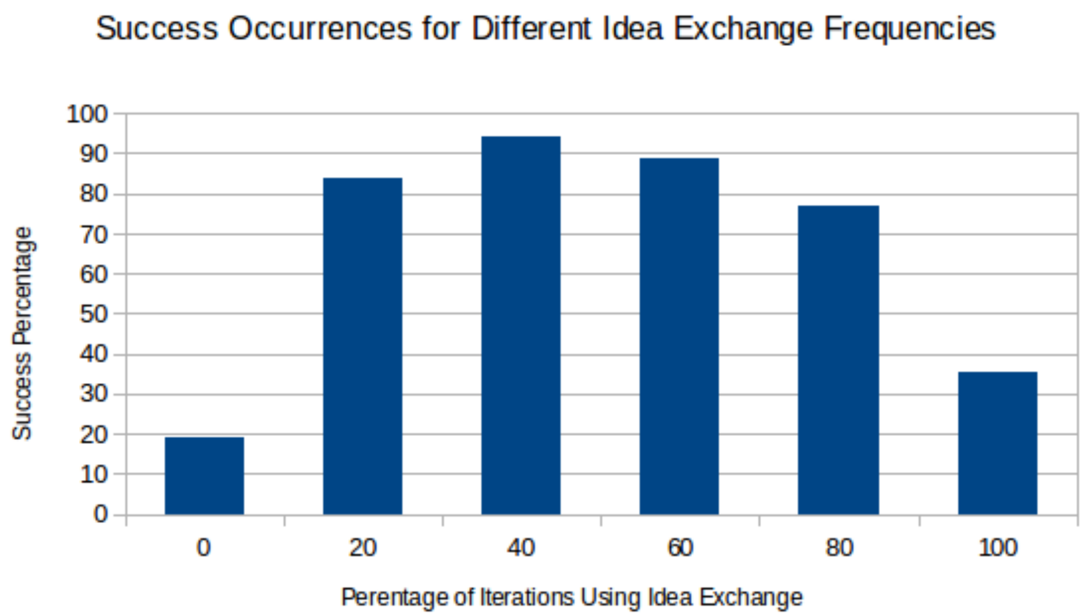

Figure 3. Success percentages for agents working to collaboratively solve problems at different idea exchange frequencies.

As previously discussed, 48 different combinations of the independent variables were run through 200 hundred simulations each. In Table 3 you can see three boxes that are outlined in bold. These correspond to the specific set of independent variables that yielded the highest problem solving success ratio. Ideal conditions for problem solving were found where the sexual percent was set to 40 percent, the agents copied ideas from random agents, and the asexual level was set to low. The analysis did not produce a split with respect to the sexual level on this branch. This branch, the bold set of three boxes, represents the global optimum conditions for the simulated collaborative problem solving trials in this study. Examining the data, one can see various branches that yield high success rates following different branching paths. These paths terminate at local optima, conditions that are good for idea exchange, but not the best.

As seen in Table 3 the criteria for which the biggest differences in success ratios appear between the first branches to second branches of the CHAID table were different for different parent first branches. The 0 percent, 40 percent, and 60 percent branches split on the basis of their asexual levels. The 20 percent sexual branch split on the basis of its sexual level. The 80 percent and 100 percent sexual branches split on the basis of the leader. For these second level splits, low asexual levels yielded better success percentages than high asexual levels, the high sexual level yielded a better success percentage than the low sexual level, and the random leaders yielded better success percentages than the best leaders.

As seen in Table 3 the criteria for the spits from the second to third branches of the CHAID table were also different for different parent second branches. Ten splits were made between these levels. Five were made on the basis of the leader variable, four of which yielded better success percentages with random leaders than with the best leaders. Three splits were made on the basis of the asexual level, all of which yielded better success percentages for the low asexual level than their high counterparts. Likewise, two splits were made on the basis of the sexual level, both of which yielded better success percentages for the low sexual level than their high counterparts.

Only four splits were made from the third level of branches to the fourth and final level of branches. Three of these splits were made on the basis of the sexual level, all of which yielded better success percentages for the low sexual level than the high counterpart. Lastly, one split was made on the basis of asexual level, yielding a better success percentage for the low asexual level than the high counterpart.

Further examination of the data draws attention to a few patterns of note. After the first set of six splits occurring on the basis of sexual percent, twenty splits occurred in the next three branching levels. All seven of the asexual level splits yielded better success rates when set to the low setting. Only one of the seven leader splits yielded a better success rate for the best leader setting, and only one of the of the six sexual level splits yielded a better success rate when set to the high setting. These two outliers appear where the sexual percent was 20. No significant sexual level or leader splits emerged where the sexual percent was 0. The theoretical implications of these findings are discussed in Chapter 5.

4.3 Summary

A total of 9,600 simulations were run in this study. The simulations accounted for all 48 combinations of the four independent variables. As discussed in Chapter 3, each trial simulated 25 agents working together to solve a complex problem for 100 creative iterations. Simulation success rates were optimized when agents exchanged ideas 40 percent of the time and individually worked on developing new problem solutions 60 percent of the time. In almost all cases the collective ability of the agents to solve problems was optimized when agents were designated to copy ideas from random collaborators as opposed to the collaborator tentatively holding the most successful solution. Likewise, in almost all cases the collective ability of agents to solve problems was optimized when agents copied small percentages of problem solutions from their collaborators. Finally, the collective intelligence of the collaborating agents was optimized when agents only changed small portions of their solutions when working individually to solve problems.

CHAPTER 5

DISCUSSION

In order to appropriately accommodate the inherent limitations of simulation based, non-empirical research, conclusions must be derived with a proper consideration of the research method's limitations and the study's theoretical framework. The purpose of this chapter is to examine the implications of this study's findings within the body of relevant theory and, in turn, to examine how these conjectures might be reinforced and expanded through future simulations and confirmed through empirical findings.

5.1 Frequency of Idea Exchange

The simulations run in this study produced data illustrating that agents most efficiently solved their designated complex problems when they utilized sexual idea exchange in balance with the solitary asexual approach to solving problems. These moderate frequencies yielded the best success ratios when they utilized sexual idea exchanges 40 percent of the time, followed closely by yields produced when sexual idea exchange was utilized 60 percent of the time. Applied to organizations, sexual exchanges represent idea exchange between collaborating individuals, while asexual represent individual level creativity. What causal phenomena could play a role in generating this pattern? The next few paragraphs explore some possible reasons for these results through a discussion of some hypothetical examples and an examination of some of the principles laid out in the literature review.

As discussed in Chapter 3, in certain circumstances genetic algorithms are superior at finding global optima than their hill climb counterparts (Mitchell, 1997). This logic would suggest that social structures that foster idea exchange or the sexual reproduction of ideas may be superior at producing globally optimal solutions to problems than counterpart social systems in which individuals tend to work in isolation. The simulations' success rates increased as the rate of idea exchange increased from 0 percent to 40 percent.

While increasing the rate of agent idea exchange in the simulation lead to improved problem solving efficiency between 0 and 40 percent, these efficiency values began to decrease as the rate of idea exchange surpassed 40 percent. As agents began to exchange ideas at rates greater than 40 percent, the collaborative network of 25 agents began to see reduced returns as a result of the system agents becoming overly interdependent. This increased rate of idea copying leads to a decrease in the variance of solutions since the agents are now more likely to zoom in on local optima that other agents may have already discovered. This may distract them from finding new, superior solutions. In a single creative iteration of the simulation the agents either mutated their old solutions or copied solutions from other agents. As agents spent less time working on their old ideas in isolation they spent less time manipulating their tentative locally optimal solutions, quickly abandoning them for ideas they borrowed from neighboring agents. With respect to asexual and sexual modes of creativity, the simulation's optimization of the system's problem solving ability parallels Marion's (2012) assertion that creativity is enhanced in systems where individuals are moderately interdependent.

Consider a hypothetical example of a four member design team working collaboratively to solve a complex engineering problem. Why might it be important for the group to properly balance their energy, time, and resource expenditures between collaboratively (sexually) working to solve the problem and individually (asexually) working to solve problem, as this study suggests? Now, suppose this design team of four members is having meeting to discuss the approaches that they have each individually developed with respect to the problem. Since these ideas were developed individually, one can be probabilistically certain that the novelty of individual ideas and the variety of total ideas within the idea space will be high. Here, these members all present their potential design solutions to each other.

Supposing that the team is examining a problem with four subcomponents and each member presents two options for each subcomponent; each member thus presents 16 possible design solutions. It is reasonable to assume that prior to the meeting a properly motivated team member could have considered the ramifications of all 16 solutions, but once all of these new ideas are presented the number of combined solutions does not increase additively to 64, rather it increases exponentially to 4,096. Obviously, no normal human being can contemplate that number of ideas in a single meeting so the discussion will likely yield itself useless after the initial ideas are passed around**.** Just like the simulated agents, the individuals in the group are only able to analyze one solution at a time. They simply see a lot of ideas, none of which they have methodically analyzed. For this reason it may be wise for the group members to part ways once they have communicated their ideas. They can then lock onto a few ideas that they are intuitively drawn to and move through proper mathematical manipulations and analytical procedures in order to analyze the combinations' potential. This can be done until the individuals' returns on continued individual contemplation begin to diminish, at which point they will need to reconvene and exchange ideas again and once again implement the newly exchanged ideas into their individual contemplations.

One of the primary problems this design group will face is created by the interaction of the design problem's four components. It is not necessarily true that each of the four design component solutions making up the ideal problem solution will still be best if the other three are changed. The problem here is that it may not be apparent that the best component solutions are the best until all four of them are combined together. As the group works to solve the problem, their design choices are path dependent (Page & Miller, 2007). As they stated, "evolutionary systems often get stuck at local optima (p. 81)." In the team members' individual time searching for problem solutions (asexual creativity) they are able to refine their ideas and search for local optima. Without doing this they may never realize when they have a potential solution. The value of their time spent collaborating exists in giving the group members the chance to compare the local optima they have isolated in hopes of finding the global optimum. This also gives them the chance to consider mixing and matching the design component solutions they have generated. Some of these components will likely be incompatible yielding low quality solutions, but others may provide them with new improved local optima and possibly even the global optimum.

5.2 Asexual Creativity Level

The CHAID tree produced through the statistical analysis of the study's 9,600 simulations split the data on the basis of asexual level seven different times. All seven of these branchings yielded superior results for the low asexual level as compared to the corresponding high asexual level. This asexual level variable represented two settings for the percentage of an agent's 25 idea digits that are mutated in a creative cycle. The high setting was likely less effective as it caused the solutions to be changed significantly enough that they were unlikely to maintain the integrity of the previous solutions. In other words, when only making small amounts of mutations the agents were able to iteratively improve on their locally optimal solutions, an ability which outweighed the ability to find new optima by making a larger number of mutations.

5.3 Leaders and Sexual Creativity Level

The CHAID tree produced through the statistical analysis of the study's 9,600 simulations split the data on the basis of sexual level six different times and the leader seven times. All but one of the sexual level splits yielded a better return at the variable's low setting. Similarly, all but one of the leader splits yielded a better return when agents were assigned to copy random agents. Both of these patterns relate to the agents' likelihood of congregating around a local optimum. Copying solutions from the best agent does this by focusing the effort of the agents on the current best solution as opposed to focusing their effort on searching for new optima. The high sexual level also led to this phenomenon of over convergence causing agents to abandon large portions of their solutions in favor of large portions of neighboring solutions, in turn forcing them to abandon other local optima before fully exploring their potential for leading to globally optimal solutions. Once again, these ideas fall in line with Marion's (2012) view that individuals in a collaborative network should be moderately interdependent.

Another pattern appears when examining the sexual level and leader outliers. A lone split occurred in the CHAID tree where agents yielded superior responses when copying the best solution, and another lone split occurred in the tree where agents yielded superior responses when set to the high sexual level. Both of these outlying patterns occurred where the agents utilized idea exchange a mere 20 percent of the time. No sexual level or leader spits occurred at the 0 percent idea exchange level, as these two variables have no bearing on systems which do not make use of sexual modes of creativity. So, these two outlying splits occurred at the lowest sexual percentage setting where any leader of sexual level splits occurred at all. Due to the low level of idea exchange here, agents would be significantly less likely to converge on suboptimal solutions, thereby the best leader and high sexual level settings must not have pushed the simulation beyond the point of over convergence, excessive interdependency.

5.4 Extrapolating to Different Degrees of Complexity

One of the largest limitations of this study is rooted in the assumption that the complexity of the simulated problem is representative of the complexity of real world problems. The non-linear nature of the problem used in the simulation yields it complex, but does not necessitate its complexity perfectly reflects that of real world problems. A couple of hypothetical examples can help deal with this issue. As described in Chapter 3, the problem-solution quality was produced as the complex interaction between an agent's 25 guess digits and their proximity to 25 ideal guess digits. Here, the solution quality yielded is dependent upon the interaction of all of the digits; therefore, no portion of the problem can be solved independently of any other.

Real world problems often have multiple complex portions that can be solved independently, yet the resulting functionality of the whole problem solution is dependent upon all of the solution portions. For example, imagine a hypothetical tribe of hunter-gatherers working to create a new hunting system, which entails the design of a new weapon and a new strategy with which the tribesmen can track animals. Here, the tribesmen have two independent complex problems to solve. Suppose that solving one makes solving the other no easier, or negligibly so. The problem emerges when the tribesmen's survival is contingent upon their ability to solve both problems. If they solve just one they will die, therefore solving one of the two problems holds no advantage over solving neither of the two problems. How might this dilemma influence the importance of maintaining idea exchange behaviors that increase their ability to collectively solve complex problems?

In the above hypothetical situation the survival of the tribe is dependent upon its ability to solve two independent problems, an issue which compounds the complexity of the dilemma. Assuming that the tribe has enough resources to work on both problems simultaneously and that the group has a probability, P, of solving each problem within the allowable time frame, then the probability of solving both problems, P_2, is represented in Equation 11.

$$P_2 = P^2 \tag{11}$$

Seeing as the probability of solving one problem, P, will always be less than one, Equation 11 illustrates that the probability of solving both problems, P_2, will always be less than P. It follows that the importance of the group's ability to solve complex problems (in this example the group's survival is contingent upon it) increases when the second problem is added. Naturally, real world situations often necessitate that individuals solve significantly more than two problems, thus the importance of being able to efficiently solve problems is compounded even further. For example, consider the number of independent complex problems involved in running an academic department or designing an internal combustion engine. Even a single failure among these multitudes of problems can cause the entire systems to collapse. Equation 12 illustrates the probability, P_N, of a group of individuals solving N problems simultaneously, assuming that the group has a probability, P, of solving each problem within the allowable time frame.

$$P_N = P^N \tag{12}$$

As N increases in Equation 12 the probability of successfully solving all problems, P_N, drops. This phenomenon is illustrated in Figures 4 and 5 showing how the probability of successfully solving all problems decreases as the number of problems increases. The graph shows an overlay of the trends for multiple hypothetical P values.

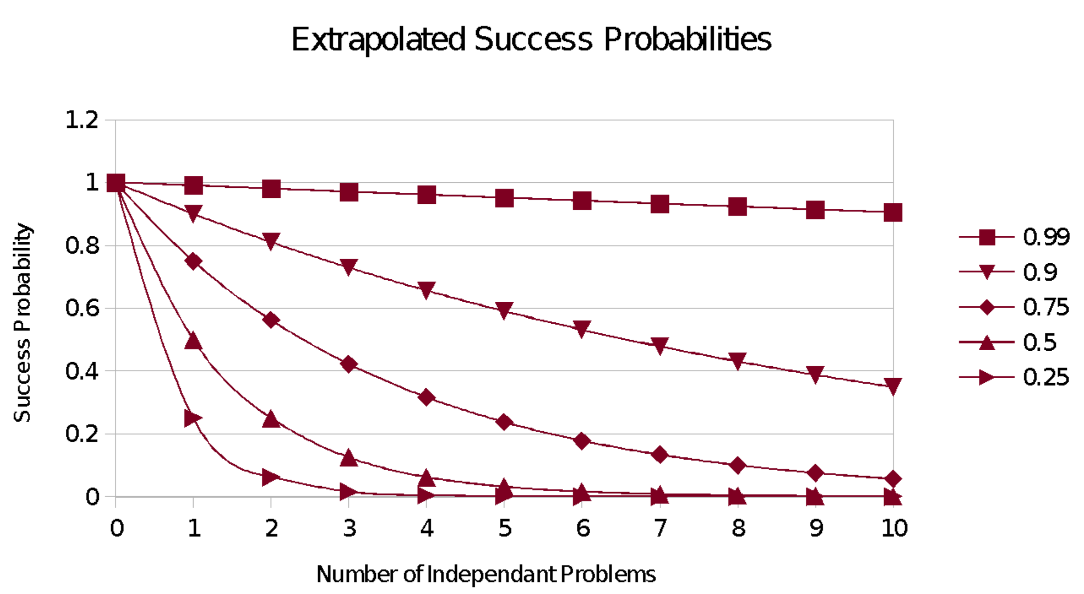

Figure 4. Probability of a group's success at collaboratively solving multiple independent problems given a probability, *P*, of solving one.

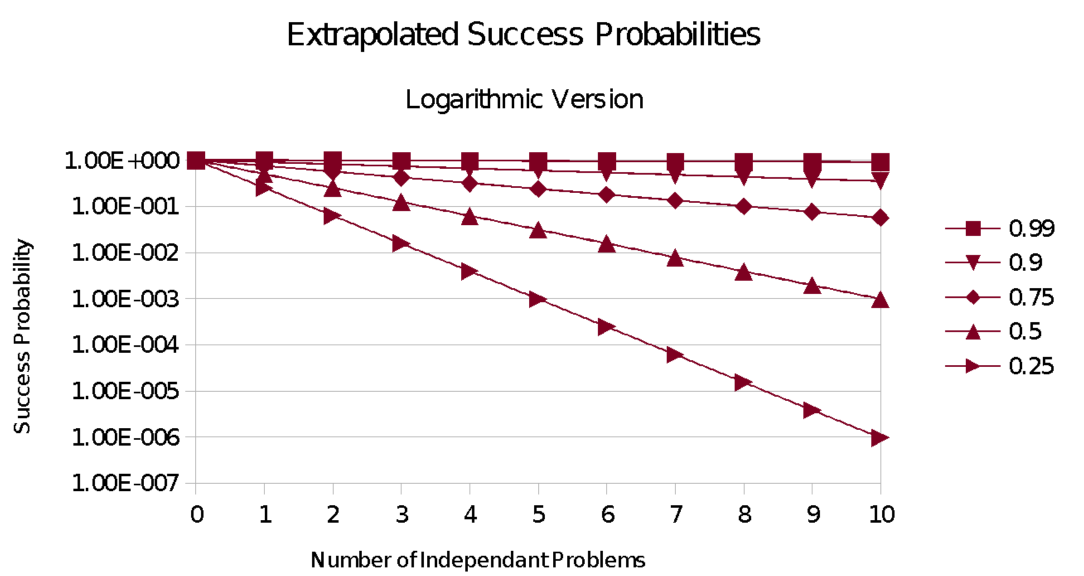

Figure 5. Probability of a group's success at collaboratively solving multiple independent problems given a probability, *P*, of solving one. The logarithmic version more appropriately shows divergence of low-end probabilities.

5.5 Future Research

This section focuses on two forms of future research that could be used to expand upon and solidify the findings of this study, future simulations and future empirical work. The discussion of future simulations examines how the simulation used in this study might be modified and expanded in order to further explore the theoretical conclusions brought forward by the work. The discussion of future empirical research focuses on the necessary steps needed to fully illustrate that the theoretical conclusions drawn from the simulations accurately reflect the behavior of actual collaborative problem solving systems.

5.5.1 Simulations

The model used in this study was made scalable in accordance with Miller and Page's (2007) recommendations. Scalable agent based models allow for the number of agents in them to be adjusted. The research in this study only examined the collaborative abilities of systems of 25 agents, so a logical next progression would be to adjust the number of agents in the simulation and observe the following effects on the agents' abilities to solve problems. In addition to exploring the effects of new independent variables further simulations could be run in order to increase the resolution of the four independent variables already under examination. Two of these variables, the asexual and sexual creativity levels were only examined at two basic levels, high and low. By incrementing these settings at smaller intervals and running an increased number of simulations the phenomena examined in this work could be analyzed in greater detail.

This work explored the influence of agents' idea copying strategies when attempting to sexually solve problems, however the study only explored two agent copying strategies, one where agents copied solutions from the agent tentatively carrying the best solution and another where agents copied solutions from an agent selected at random. This provided a solid baseline for analysis, showing the negative effects of an agent *follow the leader* strategy on optimizing collective intelligence. That said, there are other strategies that could be explored. Agents could use some form of probabilistic filter to determine with which agents they would exchange ideas or they could do so on the basis of some new parameter added into the simulation, such as a location based proximity factor or a proclivity to interact with agents whom they had previously established relationships. All simulations in this study were run for 100 iterations. Varying this number could generate even more results.

Future research could also examine different dependent variables than that of this study. This study analyzed the influence of various parameters on the success or failure of whether at least one of the 25 agents in the simulation reached the optimal problem solution within 100 iterations. It would also be possible to analyze the ability of all agents to reach an ideal solution within a given number of iterations. Other studies could examine the percentage of iterations in which the simulation agents improved upon their previous solutions or the largest number of iterations agents went without making improvements on their problem solutions (the number of iterations they spent stuck at local optima).

Another important way in which simulations could be used to further explore the nature of human collaboration is to adjust the nature of the complex problem that the agents are attempting to solve. Design requirement 3.3.1 stated that, “Real world problems are complex in nature. Solutions are dependent upon the complex interactions between different problem components,” thus, “the designed problem must be non-linear.” While this study did simulate a problem that met this design requirement it did not simulate the only possible complex non-linear problem that would fulfill this criterion. Chapter 3 discussed the concept of multidimensional fitness landscapes and explained the effectiveness of machine learning algorithms (those which parallel asexual and sexual memetic reproduction) in finding local and global peaks. Other complex problems represented by different fitness landscapes could potentially yield different responses to the independent variables manipulated in this study, thus designing new complex problems is a logical next step for this research.

The simulation used in this study used homogeneous agents which all operate identically (Page & Miller, 2007). Future research could explore the possibility of using heterogeneous agents that behave differently from each other. This opens up two primary research possibilities. The agents in this simulation could be modified to interpret problem solution quality values with a degree of variance so as to simulate agents collaboratively working to solve subjectively open problems (unlike the objectively defined problems in this study). Another possibility for modifying the agents to be heterogeneous would be to vary the behavioral patterns of the agents in the simulations. This research included simulations with 48 different combinations of four independent variables controlling agent behavior, however this variance in behavior occurred between these 48 different types of simulations rather than within them. Creating heterogeneity within a single type of simulation, such that different agents worked and collaborated differently when generating problem solutions, offers a multitude of new simulation possibilities.

5.5.2 Empirical Research

The main function of supplementing this research with empirical studies in the future would be to illustrate that the behavior of the systems modeled accurately reflect behavior of actual systems of individuals collaboratively working to solve real world problems. As discussed in length earlier in this chapter, the primary limitation of this study is a limitation inherent to non-empirical, simulations of complex social systems. For this reason, all conclusions drawn from simulations must eventually be examined under the lens of appropriate empirical testing. Two primary strategies could be used to do this empirical work. First, experimental tests could set up in order to access how various variables influence experimental groups' abilities to collaboratively solve problems. Second, analysis could be performed on collaborative systems that already exist and are already working to solve complex problems. These groups could include graduate level and undergraduate level research groups and design teams, small groups of collaborating professors, entire academic departments, larger institutions, think tanks, as well as research and development departments in various industries.

One potential way to test the influence of idea exchange on the ability of groups to collaboratively solve problems would be to run a large number of experiments in which teams of equal size would work together to solve problems and puzzles. These problems would either need to have predefined solution criteria with which groups could objectively assess solutions or a group of experts (in the field relevant to the problem) would need to be available to evaluate the solutions. Individual members of the groups would need to undergo a variety of aptitude assessments prior to the experimental trials in order to evenly distribute individual ability levels between the groups (to the greatest extent possible). By doing so, the researchers could assess the influence variables beyond group member aptitude have on the groups' abilities to solve problems.

Once the groups were established they could be given a set amount of time to generate a solution to the complex problem they were given. All groups could be given the same amount of time to solve problems, but they could be given different amounts of time to work together as opposed to working on the problems by themselves. By giving different time constraints to different groups and running a large enough number of trials on a large enough number of groups this approach would provide a tool to test some of the findings of this simulation based study.

In addition to running designed experiments to test the validity of the findings of this work researchers could also examine these findings by analyzing the behavior of collaborative organizations that already exist. A variety of innovation schemes such as Google's Eighty Twenty system are deigned to foster collaborative creativity. This Google system gives engineers and coders 20 percent of their time to exchange, foster, and develop their own ideas rather than spending time on job duties that trickle from the top down (Mediratta, 2007). The Google strategy provides an example of a system that could be used as a case study for examining the influence of collaboration on idea generation.

Seniors in Clemson University's Mechanical Engineering department have to complete capstone projects in which they work on a design project for an engineering company. They work on these projects in groups of four or five. Capstone projects are an educational practice that is also used in other engineering departments at Clemson University as well as at different universities across the country. These projects provide a large sample of collaborative groups to research. The groups could be used to help empirically identify the collaborative factors that influence the ability of systems to collectively solve problems.

Another advantage of using these groups to run studies is that universities already have individuals within their faculties with the expertise necessary to assess solution quality. With proper Institutional Review Board (IRB) approval, researchers could also obtain access to student grades in order to work to control for student aptitude. Researchers could improve their ability to control for differences between group members by providing study participants with various aptitude tests and personality inventories.

The number of methods through which researchers could study the factors that influence collaborative groups' abilities to innovate is practically limitless. Data could be collected from group meeting minutes, forms designed to track when group members collaborated, and email correspondence. Surveys could be used to develop virtual networks representing real world design groups. Group progress reports could be used to determine when design groups came up with new ideas. A number of design tools such as traditional statistical techniques, Dynamic Network Analysis, a computational method for analyzing social systems as networks of interconnected entities (Carley, 2003), or artificial neural networks (Haykin, 1994) could be used to analyze this data.

5.6 Summary

One of the primary functions of higher educational institutions is to solve problems. Academic researchers work to improve the functionality of prosthetic limbs, reduce the educational achievement gap, enhance the effectiveness of secondary school science and arts curricula, cure cancer, and increase the speed of computer processors. While these research problems lie in different academic fields, they all share two attributes. All of these problems are designed to improve our standard of living and are complex in nature. As researchers dedicate themselves to solving problems that impact the well being of people around the world, they are ethically bound to do so with the greatest efficiency possible. This ethical obligation to mankind mandates we consider all mechanisms that help us solve complex problems. In accordance with this study, the evolutionary principles that govern collective intelligence, especially those pertaining to the sexual reproduction of ideas, have significant influence on our ability to solve problems and, in turn, fulfill our ethical duties.

This study outlined a number of factors that contribute to a collaborative system's ability to collectively solve complex problems. This goal was accomplished through an extensive literature review and the development of a new computational simulation used to explore the ramifications of these ideas. The study found evidence to suggest that collective intelligence is optimized when individuals develop a balance between time spent individually working on problems and time spent exchanging ideas with others. Researchers and scholars must work to find an optimal balance between working alone and collaborating with their colleagues.

In addition to illustrating the need to balance idea exchange with solitary idea generation, the simulation provided evidence that individuals should avoid blindly following current, cutting edge solutions to problems. While these solutions certainly have significant societal value, premature adoption by all creative thinkers may cause collective systems to reduce their chances of finding new solutions that are even better than the current optima. Here, researchers should be provided with the freedom to explore new ideas that lay in opposition to traditional perspectives before pushing them to use proven methods and theories. When properly balanced, the cost of having researchers waste time on a few dead ends will be outweighed by the value of the new ideas this approach is likely to produce. Universities, academic departments, thank tanks, and educational policy makers can increase the problem solving efficiency of various institutions by promoting conditions which provide researchers in all fields with extensive networks in which they can exchange ideas and by promoting conditions that moderate the interdependency of individuals.

APPENDICES

Appendix A

SPSS Results													
Node	Fail N	F.%	Success N	S.%	N	%	Predicted Category	Parent Node	Primary Ind. Variable	Sig. a	Chi-Square	df	Split Values
0	3238	33.70%	6362	66.30 %	9600	100.00 %	1	n/a	n/a	n/a	n/a	n/a	n/a
1	1298	81.10%	302	18.90 %	1600	16.70%	0	0	S.P.	0	3502.242	5	<= .0
2	262	16.40%	1338	83.60 %	1600	16.70%	1	0	S.P.	0	3502.242	5	(.0, 20.0]
3	94	5.90%	1506	94.10 %	1600	16.70%	1	0	S.P.	0	3502.242	5	(20.0, 40.0]
4	182	11.40%	1418	88.60 %	1600	16.70%	1	0	S.P.	0	3502.242	5	(40.0, 60.0]
5	368	23.00%	1232	77.00 %	1600	16.70%	1	0	S.P.	0	3502.242	5	(60.0, 80.0]
6	1034	64.60%	566	35.40 %	1600	16.70%	0	0	S.P.	0	3502.242	5	> 80.0
7	585	73.10%	215	26.90 %	800	8.30%	0	1	A.L.	0	66.874	1	1
8	713	89.10%	87	10.90 %	800	8.30%	0	1	A.L.	0	66.874	1	2
9	175	21.90%	625	78.10 %	800	8.30%	1	2	S.L.	0	35.345	1	1
10	87	10.90%	713	89.10 %	800	8.30%	1	2	S.L.	0	35.345	1	2
11	13	1.60%	787	98.40 %	800	8.30%	1	3	A.L.	0	52.262	1	1
12	81	10.10%	719	89.90 %	800	8.30%	1	3	A.L.	0	52.262	1	2
13	44	5.50%	756	94.50 %	800	8.30%	1	4	A.L.	0	54.781	1	1
14	138	17.30%	662	82.80 %	800	8.30%	1	4	A.L.	0	54.781	1	2
15	89	11.10%	711	88.90 %	800	8.30%	1	5	L.	0	127.4	1	1
16	279	34.90%	521	65.10 %	800	8.30%	1	5	L.	0	127.4	1	2
17	348	43.50%	452	56.50 %	800	8.30%	1	6	L.	0	312.332	1	1
18	686	85.80%	114	14.30 %	800	8.30%	0	6	L.	0	312.332	1	2
19	123	30.80%	277	69.30 %	400	4.20%	1	9	L.	0	36.871	1	1
20	52	13.00%	348	87.00 %	400	4.20%	1	9	L.	0	36.871	1	2
21	29	7.30%	371	92.80 %	400	4.20%	1	10	A.L.	0.0 01	10.846	1	1
22	58	14.50%	342	85.50 %	400	4.20%	1	10	A.L.	0.0 01	10.846	1	2
23	1	0.30%	399	99.80 %	400	4.20%	1	11	L.	0.0 02	9.461	1	1
24	12	3.00%	388	97.00	400	4.20%	1	11	L.	0.0	9.461	1	2

				%						02			
25	22	5.50%	378	94.50 %	400	4.20%	1	12	L.	0	18.805	1	1
26	59	14.80%	341	85.30 %	400	4.20%	1	12	L.	0	18.805	1	2
27	12	3.00%	388	97.00 %	400	4.20%	1	13	L.	0.0 02	9.62	1	1
28	32	8.00%	368	92.00 %	400	4.20%	1	13	L.	0.0 02	9.62	1	2
29	33	8.30%	367	91.80 %	400	4.20%	1	14	L.	0	45.396	1	1
30	105	26.30%	295	73.80 %	400	4.20%	1	14	L.	0	45.396	1	2
31	33	8.30%	367	91.80 %	400	4.20%	1	15	A.L.	0.0 1	6.688	1	1
32	56	14.00%	344	86.00 %	400	4.20%	1	15	A.L.	0.0 1	6.688	1	2
33	102	25.50%	298	74.50 %	400	4.20%	1	16	A.L.	0	30.958	1	1
34	177	44.30%	223	55.80 %	400	4.20%	1	16	A.L.	0	30.958	1	2
35	144	36.00%	256	64.00 %	400	4.20%	1	17	S.L.	0	18.309	1	1
36	204	51.00%	196	49.00 %	400	4.20%	0	17	S.L.	0	18.309	1	2
37	302	75.50%	98	24.50 %	400	4.20%	0	18	S.L.	0	68.784	1	1
38	384	96.00%	16	4.00%	400	4.20%	0	18	S.L.	0	68.784	1	2
39	45	22.50%	155	77.50 %	200	2.10%	1	19	A.L.	0	12.785	1	1
40	78	39.00%	122	61.00 %	200	2.10%	1	19	A.L.	0	12.785	1	2
41	40	20.00%	160	80.00 %	200	2.10%	1	30	S.L.	0.0 04	8.071	1	1
42	65	32.50%	135	67.50 %	200	2.10%	1	30	S.L.	0.0 04	8.071	1	2
43	38	19.00%	162	81.00 %	200	2.10%	1	33	S.L.	0.0 03	8.896	1	1
44	64	32.00%	136	68.00 %	200	2.10%	1	33	S.L.	0.0 03	8.896	1	2
45	72	36.00%	128	64.00 %	200	2.10%	1	34	S.L.	0.0 01	11.036	1	1
46	105	52.50%	95	47.50 %	200	2.10%	0	34	S.L.	0.0 01	11.036	1	2

Growing Method: CHAID
Dependent Variable: Success
a Bonferroni adjusted

Appendix B

%GNU Octave Code
%Noah Welsh
%A Computational Model of Memetic Evolution

% This creates a matrix storing the simulation results. Set first number in parentheses to the total number of simulations (Number of simulations * Number of iterations in each).

```
Best_100b= zeros(196000,7);
```

%Imbedded loops walk the code through each combination of independent variables.

```
for lead=1:2
        for ga=1:2
                for hc=1:2
                        for ratio =1:6
```

%To run *N* trials for each combination of independent variables adjust the line below to for trials =1:N

```
                                for trials =1:200
                                        if ratio ==1
                                                algorithmratio=0;
                                        elseif ratio ==2
                                                algorithmratio=5;
                                        elseif ratio ==3
                                                algorithmratio=9;
                                        elseif ratio ==4
                                                algorithmratio=13;
                                        elseif ratio ==5
                                                algorithmratio=17;
                                        else
                                                algorithmratio=21;
                                        end
                                        Coeff =zeros (25,25,'uint32');
```

%Randomly generates weighting coefficients.

```
                                        for i=1:25
                                                for j=1:25
                                                        Coeff(i,j)=randi(10)-1;
                                                end
                                        end
```

```
for i=1:25
        output(i)=99999999;
end
```

% Next, we randomly assign a 25 (adjust this appropriately) digit code representing external reality. As we guess at the nature of reality, the closer our estimates get to the true nature of reality, the better our results get. It is important to remember these results are not binary. When designing a spear, car, or higher education financial model, there are practically infinite possibilities. These design ideas are not right or wrong. They lie along a spectrum from better to worse. My model assumes that low return values are better than high return values. Here, humans use induction. They guess at how something will work, use creativity (random mutation of old ideas), and try again. If their result improve they will keep their new ideas and mutate them. If the mutations make the returned outputs worse, then the agents revert to old designs and repeat. %

% Initial guesses are randomly assigned. This essentially illustrates the idea that the first humans had to start from scratch when developing ideas on any real world problems. Cave people had no preconceptions about how to approach phenomena of practically infinite complexity infinite complexity.

```
for i=1:25
        Truth(i)= randi(10)-1;
        for j=1:25
                Guess(i,j)= randi(10)-1;
                Guess2(i,j)=Guess(i,j);
        end
```

%Uncoment the line below to see the actual problem solution values.

%printf("%f ", Truth(i));

```
end
```

% HYPOTHESIS: This individual creativity represents the creativity of animals, and the creativity of humans prior to the development of language (the ability to exchange ideas between agents).

% NOTE: Language is not a binary phenomenon. Language has evolved so that idea exchange mechanisms have been developing for millions of years. Here, we see the biological and anthropological departure from instinct (genetically rooted ideas) to learned behaviors (ideas which are nurtured). */

% Initial output must be set to high value so that my model does not have early humans assume they have perfectly figured the nature of reality

```
BestOutput=99999999;
Avg=1;

%To adjust the simulation trials to run for X iterations change the linee below to for
f=1:X
for f=1:100
	total=total+1;

	for i=1:25
		Noutput(i)=0;
	end
	for e=1:25
		for i=1:25
			j=i+1;
			while j<26

Noutput(e)=Noutput(e)+power(Coeff(i,j)*(Guess(e,i)*Guess(e,j)-Truth(i)*Truth(j)),2);
				j=j+1;
			end
		end
	end

%This compares new solution values to the values of previous iterations.

	for i=1:25
		if Noutput(i)<output(i)
			for j=1:25

	Guess2(i,j)=Guess(i,j);
			end
			output(i)=Noutput(i);
		else
			for j=1:25

	Guess(i,j)=Guess2(i,j);
			end
			Noutput(i)=output(i);
		end
		if output(i)<BestOutput
			Best=i;
			BestOutput=output(i);
		end
	end
```

```
BestOut(f)=BestOutput;
Total=0;
for z=1:25
Total=Total+Noutput(z);
end
```

%Calculates the average solution value.

```
Avg(f)=Total/25;
```

%The lines below store the simulation results in a matrix.

```
Best_100b(total,1) = f;
Best_100b(total,2) = lead;
Best_100b(total,3) = hc;
Best_100b(total,4) =ga;
Best_100b(total,5) = ratio;
Best_100b(total,6) = BestOutput;
Best_100b(total,7) = Avg(f);
printf("%i, %i,%i, %i\n",BestOutput, Avg(f),
f, total);

for z=1:25
```

%The code below adjust the asexual and sexual level settings

```
if hc==1
        hcsize=randi(7);
else
        hcsize=randi(18);
end
if ga ==1
        gasize=randi(7);
else
        gasize=randi(18);
end
```

%The line below reads, “if randi(20)<algorithmratio” The probability of an agent utilizing sexual modes of creativity in a given iteration, P = (algorithmratio – 1)/20

```
if randi(20)<algorithmratio
        for num1 =1:gasize
                if lead==1
```

```
        copy=randi(25);

        choice=randi(25);
                                                                    Guess(z,copy)
=Guess2(choice,copy);
                                                              else
copy=randi(25);
                                                                    Guess(z,copy)
=Guess2(Best,copy);
                                                              end
                                                        end
                                                  else
                                                        for num2 =1:hcsize
                                                              Guess(z,randi(25))
=randi(10)-1;
                                                        end
                                                  end
                                            end
                                      end
                                end
                          end
                    end
        end
end
```

%Uncomment the lines below to output a graph of the best solution as a function of the iteration number. Change “BestOut” to “Avg” to obtain a graph of average solution as a function of the number of iterations. Generally speaking, this is only useful when a single simulation trial is being run.

```
% semilogy(BestOut)
% xlabel('Creative Iterations')
% ylabel('Problem Solution Proximity to Ideal Value')
% title('Creative Cycles Required for Collaborating Agents to Reach Ideal Solutions')
```

REFERENCES

Anderson, P. W. (1972). More is different. *Science (177)*, 393-96.

Audretsch, D. B. & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review (86),* 630-640.

Austin, A. E. (1990). Faculty cultures, faculty values. In W. G. Tierney (Ed.), Assessing academic climates and cultures (pp. 61-74). San Francisco: Jossey-Bass.

Baker, R. R. (1966). *Sperm wars: Infidelity, sexual conflict and other bedroom battles.* London: Fourth Estate.

Baldwin, J. M. (1896). A new factor in evolution. *American Naturalist,* 30, 441-51, 536-53.

Baldwin, J. M. (1909). *Darwin and the humanities*, Baltimore, MD: Review Publishing.

Berger, J. (2013). *Contagious: Why things catch on.* New York: Simon & Schuster.

Bonabeau, E. (2002). *Agent-Based modeling: Methods and techniques for simulating human systems.* Proceedings of the National Academy of Sciences of the United States of America, 99(10): 7280–7287.

Bonabeau, E. & Meyer, C. (2001). Swarm intelligence: A whole new way to think about business, *Harvard Business Review, 79*(5), 106- 114.

Boyatzis, R. E. (2006). An overview of intentional change from a complexity perspective. *Journal of Management Development 25*(7), 607-623.

Boyer, E. (1990). *Scholarship reconsidered: Priorities of the professoriate.* Princeton, NJ: Princeton University Press.

Boyer, E. (1997). A community of scholars. In E. Boyer (Ed.), *Selected speeches, 1979-1995* (pp. 69-80). Paper delivered at the Emory Symposium, Atlanta, GA, April 14, 1994. San Francisco: Jossey-Bass.

Buss, D. M. (1994). *The evolution of desire: Strategies for human mating,* New York: Basic Books.

Campbell, D. T. (1960). *Blind variation and selective retention in creative though as in other knowledge processes.* Psychological Review, 67, 380-400.

Campbell, D. T. (1965). Variation and selective retention in sociocultural evolution. In H. R. Barringer. G. L. Blanksten and R. W. Mack (Eds.), *Social change in developing areas: A reinterpretation of evolutionary theory* (pp. 19-49). Cambridge, MA: Schenkman.

Campbell, N. A., & Reece, J. B. (2011). *Campbell biology*. San Francisco, CA: Benjamin Cummings.

Carley, K. (2003). Dynamic network analysis. In R. Breiger & K. Carley (Eds.), the summary of the National Research Council Workshop on Social Network Modeling and Analysis, National Research Council.

Cavalli-Sforza, L. L. & Feldman, M. W. (1981). *Cultural evolution and transmission: A quantitative approach.* Princeton, NJ: Princeton University Press.

Cheng, Y. & Van de Ven, A. (1996). Learning the innovation journey: Order out of chaos. *Organization Science, 7,* 593-614.

Christiansen, J. K. & Varnes, C. J. (2008). Making decisions on innovations: Meetings or networks? *Creativity and Innovation Management 16*(3). 282-298.

Clark, B. (1983). *The higher education system: Academic organization in cross-national perspective.* Berkeley: University of California Press.

Damrosch, D. (1995). We scholars: Changing the culture of the university. Cambridge: Harvard University Press.

Darwin, C. (1859). *On the origins of species by natural selection.* London: Murray.

Darwin, E. (1794). *Zoonomia.* (Vol. 1). London: Johnson.

Dawkins, R. (1978). *The selfish gene.* Oxford University Press.

Deacon, T. (1997). *The symbolic species: The co-evolution of language and the human brain.* London: Penguin.

Delius, J. (1989). Of mind memes and brain bugs, a natural history of culture. In W. A. Koch (Ed.), *The Nature of Culture*, (pp 26-79). Bochum, Germany: Bochum Publications.

Dennett, D. C. (1991). *Consciousness explained.* Boston: Little, Brown and Co.

Dennett, D. C. (1994). *Darwin's dangerous idea.* London: Penguin.

Dennett, D. C. (2004). *Freedom evolves.* Penguin (Non-Classics).

Desai N. S. & Miikkulainen R. (2000). Neuro-evolution and natural deduction. In *Proceedings of The First Symposium on Combinations of Evolutionary Computation and Neural Networks*, 64-69, Piscataway, NJ.

DeSalle, R. & Tattersall, I. (2012). *The brain.* New Haven, CT: Yale University Press.

Diggins, J. P. (1994). *The promise of pragmatism: Modernism and the crisis of knowledge and authority.* Chicago, IL: University of Chicago Press.

Dill, D. D. (1991). The management of academic culture: Notes on the management of meaning and social integration. In J. L. Bess (Ed.), Foundations of American higher education (pp. 567-579). Needhain Heights, MA: Ginn Press.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition.* Cambridge, MA: Harvard University Press.

Donald, M. (1993). *Précis of* Origins of the modern mind: Three stages in the evolution of culture and cognition. *Behavioral and Brain Sciences,* 16, 737-91.

Draper, N. R. & Smith, H. (1998). *Applied regression analysis.* John Wiley.

Dunbar, R. (1996). *Grooming, gossip and the evolution of language.* London: Faber and Faber.

Duhram, W. H. (1991). *Coevolution: Genes, culture and human diversity.* Stanford, CA: Stanford Press.

Eaton, J. W. (2013). GNU Octave (Version 3.6.4) [Software]. Available from http://www.gnu.org/software/octave/doc/interpreter

Ethiraj, S. K. & Levinthal, D. (2011). Modularity and Innovation in Complex Systems. *Management Science 50*(2), 159.

Fisher, R. A. (1930). *The genetic theory of natural selection.* Oxford University Press.

Gentle, J. E. (2004). *Random number generation and Monte Carlo methods*. Springer.

Gilbert, D. T. (2007). *Stumbling on happiness.* New York: Vintage Books.

Gilder, G. (2000). *Telecosm: How infinite bandwidth will revolutionize our world.* New York: Press.

Gödel, K. (1962). *On formally undecidable propositions of Principia Mathematica and related systems*. Edinburgh: Oliver & Boyd.

Goldstein, J. (1999), Emergence as a construct: History and issues, *Emergence: Complexity and Organization (1)*1, 49–72.

Grant, G. (1990). Memetic Lexicon. http://pepmcl.vub.ac.be/*memes.html

Hargadon, B. & Bechky, B. A. (2006). When collections of creatives become creative collectives: A field study of problem solving at work. *Organization Science 17*(4), 484-500.

Haykin, S. (1994). *Neural network: A comprehnsive foundation.* New York: Macmillan College Publishing Company.

Holland, J. H. (1975). *Adaptation in natural and artificial systems.* Ann Arbor, MI: University of Michigan Press.

Holland, J. H. (1994). *Hidden order: How adaptation builds complexity.* New York: Basic Books.

Hull, D. L. (1992). The naked meme. In H. C. Plotkin (Eds.), *Learning, Development and Culture,* (pp. 273-327). London: Wiley.

IBM Corp. (2013). IBM SPSS Statistics for Windows (Version 22.0) [Software]. Armonk, NY: IBM Corp.

Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review (76).* 984-1001.

Jones, R. H. (2013). Analysis & the fullness of reality: An introduction to reductionism & emergence. Jackson Square Books.

Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data, *Applied Statistics*, *(29)*2, 119–127.

Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution.* Oxford University Press.

Kauffman, S. A. (1995). *At home in the universe: The search for laws of complexity.* Oxford University Press.

Kerr, C. (1982). *The uses of the university.* Cambridge, MA: Harvard University Press.

Kurzweil, R. (2006). *The singularity is near: When humans transcend biology.* Gerald Duckworth & Co Ltd.

Latucca, L. (2002). Learning interdisciplinarity: Sociocultural perspectives on academic work. *The Journal of Higher Education. (73)*6, 711-739.

Levy, P. (1997) *Collective intelligence: Mankind's emerging world in cyberspace.* Basic Books.

Lynch, A. (1991). Thought contagion as abstract evolution. *Journal of Ideas,* 2, 3-10.

Marion, R. (2012). *Leadership of creativity: entity-based, relational, and complexity perspectives.* In the Handbook of Organizational Creativity.

Mathworks. (2013). *Matlab primer.* Retrieved from http://www.mathworks.com/help/pdf_doc/matlab/getstart.pdf.

McConnell, S. (1993). *Code complete: A practical handbook of software construction.* Redmond, Washington: Microsoft Press.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology,* 27, 415-444.

Mediratta, B. (2007, October 21). The Google way: Giving engineers room. *The New York Times.* Retrieved from http://www.nytimes.com/2007/10/21/jobs/21pre.html.

Miller, J. H. & Page, S. E. (2007). *Complex adaptive systems an introduction to computational models of social life.* Princeton, NJ: Princeton University Press.

Mitchell, M. (1997). *Introduction to genetic algorithms.* Cambridge, MA: MIT Press.

Mohri, M., Rostamizadeh, A., Talwalkar, A., & Ebrary, I. (2012). *Foundations of machine learning*. Cambridge, MA: MIT Press.

Nelson, B., Barreno, M., Jack, F., Anthony C., Joseph, D. , Rubinstein, B. I. P., Saini, U., Sutton, C., Tygar, J. D., & Xia, K. (2008, April). *Exploiting machine learning to subvert your spam.* Proceedings of the First Workshop on Large-scale Exploits and Emerging Threats. San Francisco, CA.

Newman, M. J. (2003). The structure and function of complex networks. *SIAM Review,* 4*,* 167-256.

Orlando, M. J. & Verba, M. (2005). Do only big cities innovate? Technological maturity and the location of innovation. *Economic Review (90)*2, 31-57.

Page, S. E. (2006). Essay: Path dependence. *Quarterly Journal of Political Science,* 1. 67-87.

Perry-Smith, J. E. & Shalley, C. E. (2003). The social side of creativity: A static and dynamic social network perspective. *Academy of Management Review 28*(1), 89-106.

Pinker, S. (1994). *The language instinct.* New York: Morrow.

Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences,* 13, 707-84. (with commentaries by others).

Popper, K. R. (1979). *Objective knowledge: An evolutionary approach*. Oxford: Clarendon Press.

Ridley, M. (1993). *The red queen: Sex and the evolution of human nature.* London: Viking.

Ridley, M. (2010). *The rational optimist: How prosperity evolves.* London: Fourth Estate.

Ridley, M. (2010). Matt Ridley: When ideas have sex. [Video File] Retrieved from http://www.ted.com/talks/matt_ridley_when_ideas_have_sex.html

Rogers, E. M. (2003). *Diffusion of innovations.* New York, Simon and Schuster, Inc.

Salakhutdinov, R. & Mnih, A. (2008). Probabilistic matrix factorization. *Advances in Neural Information Processing (20),* Cambridge, MA: MIT Press.

Schut, M. C. (2010). On model design for simulation of collective intelligence, *Information Sciences 180*, 132-155.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations,* (ed. A. Skinner), New York: Pelican Books.

Speel, H. C. (1995). *On a conceptual framework for cultural evolution.* Paper presented at the symposium 'Einstein meets Magritte' Free University of Brussels, June.

StatSoft, Inc. (2013). Popular Decision Tree: CHAID Analysis, Automatic Interaction Detection. In the *Electronic Statistics Textbook*. Retrieved from https://www.statsoft.com/Textbook/CHAID-Analysis.

Surweiki, J. (1995). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies and* nations. New York: Anchor Books.

Symons, D. (1979). *The evolution of human sexuality.* New York: Oxford University Press.

Tarde, G. (1969). *The laws of imitation.* (E. C. Parsons, Trans.). New York: Holt. (Original work published in 1903).

Tierney, W. G., & Rhoads, R. A. (1994). Faculty socialization as cultural process: A mirror of institutional commitment. Washington DC: The George Washington University. (ASHE-ERIC Higher Education Report No. 93-6)

Toth, N. & Schick, K. (1993). Early stone industries and inferences regarding language and cognition. In K. Gibson and T. Ingold (Eds.), *Tools, Language and Cognition in Human Evolution*, (pp. 346-362). Cambridge University Press.

Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual Selection and the Descent of Man,* (pp 139-179). Chicago, IL: Aldine de Gruyter.

Tsai, W. (2001). Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management Journal 44*(5), 996-1004.

Uhl-Bien, M., Marion, R., & McKelvey, B. (2007). Complexity leadership theory: Shifting leadership from the industrial age to the knowledge era. *The Leadership Quarterly 18*(4), 298-318.

Vygotsky, L. S. (1981). The genesis of higher mental functions. In J. V. Wertsch (Ed.), The concept of activity in Soviet psychology (pp. 134–143). Armonk, NY: M. E..

Williams, G. C. (1966). *Adaptation and natural selection.* Princeton, NJ: Princeton University Press.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N. and Malone, T. (2010), Evidence for a collective intelligence factor in the performance of human groups, *Science 330*, 686-688.

Wright, R. (1994). *The moral animal.* New York: Pantheon.