2011-01-01

The Exploratory Value of Agent-Based Models in Social Science

Ricardo Andress Rivera
University of Texas at El Paso, resol1@hotmail.com

THE EXPLORATORY VALUE OF AGENT-BASED MODELS IN SOCIAL SCIENCES

RICARDO ANDRES RIVERA

Department of Philosophy

APPROVED:

John Symons, Ph.D., Chair

Josiah Heyman, Ph.D.

Peter Robinson, Ph.D.

Benjamin C. Flores, Ph.D.
Interim Dean of the Graduate School

Dedication

This is dedicated to my big boy Elijah and to the younger ones in my family. I did it, this proves that you can all do it as well. It's your turn!

THE EXPLORATORY VALUE OF AGENT-BASED MODELS IN SOCIAL SCIENCES

By

RICARDO ANDRES RIVERA, B.A

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

Department of Philosophy

THE UNIVERSITY OF TEXAS AT EL PASO

December 2011

Acknowledgements

I’d like to acknowledge Shortysan and my parents Jose and Irene, without their continued support this wouldn’t have been possible.

Table of Contents

Chapter 1 Introduction

Social scientists have, what seems like, countless methods of investigation at their disposal. For a comprehensive introduction to many of the methods of social science one might review a book titled *Handbook of Methods in Cultural Anthropology* by H. Russell Bernard, an anthropologist who has written extensively on the methods of anthropological investigation. He states that “Whatever our epistemological differences, however, the actual methods by which we collect and analyze our data, these methods belong to everyone across the social science”(14). Although the book offers an extensive and detailed discussion on the methods used in anthropology and consequently social science in general, there is one method that is not discussed in the book. This method is relatively new and intends to incorporate the explosive progression of computational power that is now at most scientists’ disposal. This new method is a way to potentially investigate some of the most complex social and dynamical systems that have otherwise eluded other methods of scientific investigation. The method has distinctive features that allow for new approaches of modeling to be considered. This method will require and promote interdisciplinary work. The method is called agent-based models.

There are features of agent-based models that should make them especially interesting to social scientists. For example, the components or agents, as described above, react to and are affected by the agents that surround them. The environment or landscape that the agents are located in is constantly changing. Agents are subject to learning or mutation and selection. These are just a few of the features of agent-based models that can easily be understood as analogous features of societies and cultures[1].

[1] These features are further described throughout section 2.1. A description of how these features were applied to the Artificial Anasazi simulation is given as well.

Complex systems have eluded traditional methods of investigation because the system is not available or feasible to study, or the mathematics involved are too difficult. Agent-based models have offered some solutions to some of the obstacles that have made most complex systems difficult to study. Agent-based models are by no means limited to the social sciences. For example, in physics, cellular automata were used to model ferromagnetism. Agent-based models have also modeled the flocking behavior of birds, fish and other species, as well as termites arranging wood debris. Business alliances for promoting technical standards have also been modeled with the use of agent-based models(Humphreys, 2004,130). In social sciences, agent-based models will become a sophisticated tool that will allow for investigation into some of the complex and dynamical adaptations and trends of cultures and societies from the past present and future.

Agent-based models will facilitate the interdisciplinary work that is necessary to investigate the complex systems. There is a growing idea amongst researchers and scientists that involves the realization that the complex, social phenomena that we are interested in understanding cannot be reduced and divided nicely into parts that can be studied by a single discipline. Mitchell correctly asserts that as "the lines between disciplines begin to blur, the context of scientific discourse also gets fuzzier" (xii).

There is a continuing recognition that in order to inquire and understand complex systems, new ideas and new approaches are needed in order to make sense of the highly complex, complicated connected systems that increasingly affect us as humans. Agent-based models will be necessary in progressing our scientific knowledge and will be a catalyst for the interdisciplinary work that will be necessary in the exploration of complex systems.

Computational models and computer science have changed the epistemology of science. An accurate observation by Mitchell states that "the traditional division of science into theory and experiment has been complemented by a new category: computer simulations"(209). This "new category" and the interdisciplinary nature of agent-based models will force philosophers to concern themselves with non-traditional problems within philosophy. Philosophers will need to become familiar with issues of application in order for their ideas and comments about computational models to be valued. A talented philosopher was once asked if computational science and computer simulations were forcing philosophers out of the "arm-chair". The philosopher replied "no, not if the computer is attached to that arm-chair".

In order to understand the complex systems that interest philosophers, they will need to be part of the fundamental building blocks of the science. Philosophers will need to reevaluate their role in the philosophy of science and take on questions and problems that have not traditionally been part of most philosophers' concerns.

This thesis will introduce social scientists to a new method of inquiry and investigation, while encouraging philosophers to move from their traditional concerns of what is possible in theory to concerns of what is possible in application. Part of these building blocks will include a new vocabulary and this thesis will add to the new vocabulary through a term that is referred to as "exploratory value."

Chapter 2 is meant to aid the reader in two ways: first, the chapter will serve as an introduction to agent-based models in social science in general; second, the chapter describes a case study where an agent-based model was applied to investigate the disappearance of the ancient Pueblo people. Although, there are numerous issues introduced that are important to philosophers and social scientists, Chapter 2 will give the reader some of the general ideas and concepts behind agent-based models and the approach of a group of researchers called generative social scientists, who have applied the models in social science. Generative social scientists' claims are reviewed because they value simulations. The issue concerning the value of a simulation becomes a central issue in this thesis. Chapter 3 suggests that a more humble claim about the results of simulations must be given and supported.

Chapter 3 introduces the distinctive value of computational science and computer simulations. Exploratory value of computational models will be explained the value of computational models within social sciences will be asserted. Chapter 4 also argues that social scientists should not assume that agent-based models will give them the causal explanations to all social phenomena, but instead argues that philosophers should look for the exploratory value of their simulations.

Winsberg's ideas on how a simulation gains reliability are used to build up the concept of exploratory value. In addition, Humphreys's ideas of "extending ourselves" are reviewed and and it is argued that these ideas are a feature of the exploratory value of a simulation. Grüne-Yanoff is reviewed in order to argue that researchers will understand that their simulations should be used to gain insight into the phenomena under investigations. Subsequently, Grüne-Yanoff 's ideas of evidential support for simulations, potential functional use of simulations and the quality of simulation will illustrate how those ideas coincide with the concept of exploratory value.

In chapter 4, the concept of exploratory value is examined and a brief discussion that argues in favor of philosophers' use of computational science and computer simulations and their application is developed. This last idea implies that the building, testing, running and evaluation of agent-based models will be an interdisciplinary venture, which in turn will lead to philosophers re-evaluating the epistemology of science and their traditional roles in the philosophy of science.

Chapter 2: Generative Social Science

The goal of generative social science is to explain macroscopic phenomena and societal regularities. Epstein claims that, "to the generativist, explaining the emergence of macroscopic societal regularities, such as norms or price equilibrium, requires that one answer the following question: How could the decentralized local interactions of heterogeneous autonomous agents generate the given regularity?"(5). The generativists believes that they cant "situate an initial population of autonomous heterogeneous agents in the relevant spatial environment; allow them to interact according to simple local rules and there by generate—or "grow", the macroscopic phenomenon from bottom up". One of the driving questions in this thesis is to understand what value, agent-based models have, if they fail "to explain macroscopic phenomena".

A study used by Epstein and other generativists, that involved agent-based models and claimed to have empirical value in social science was referred to as "The Artificial Anasazi Project". One of the goals of the Artificial Anasazi Project was to "grow" or reconstruct 500 years of Ancient Pueblo Peoples history, starting in 800 AD and ending in 1300 AD. Another goal was to reconstruct the physical environment, using dendroclimatalogical and other data of the Long House Valley, a small region in the northeastern Arizona (12). Epstein states that the "basic issue is whether environmental (i.e. subsistence) factors alone can account for their sudden disappearance"(13). He questions whether other factors such as, property rights, clan relationship, conflict, disease- have to be added to generate the true history"(13). Another objective was to work with anthropologists in order to find ethnographically plausible rules of agent behavior that would generate the "true history" (13). The idea that ethnographies can be used to find plausible agent behavior becomes an issue in chapter 4, where Grüne-Yanoff's claims that the ethnographies used by Epstein, are too vague and cannot help in developing rules of behavior.

Epstein and his colleagues refer to initial rules of behavior and environmental descriptions, as a micro-specification (7). Epstein best sums up the model when he states that the challenge was "to place artificial Anasazi where the true ones were in 800 AD and see if-under the postulated rules the simulated evolution matches the true one" (13). The modelers concluded that the model of the simulated Anasazi history "closely reproduced the main features of its actual history, including population ebb and flow, changing spatial settlement patterns and eventual rapid decline"(Epstein, 9).

2.1 Characteristics of Agent-Based Models Applied

According to Epstein and other generativists, agent based computational models have distinct characteristics that make it possible to "grow" or generate macroscopic phenomenon. In Epstein's view, the central contribution of the agent-based model is that it will facilitate "*generative* explanation"(51). The characteristics in agent-based models that will facilitate explanations include; heterogeneity, autonomy, explicit space, local interactions, bounded rationality and non-equilibrium dynamics. The conceptual importance of these characteristics is that they are analogous to the kinds of interactions performed by agents in many past and present societies. The use of agent-based simulations is likely to become a valuable tool in the social sciences because often the systems of interest, such as economical trends and cultural trends, in social sciences are systems that are quite complex and are rapidly changing. In order for social scientists to be better study these kinds of systems of interest, they will need to be aided in their investigation by the calculative power in computational methods. Computational methods will allow social scientists to track and represent individual agents reacting to other agents and their environments, which can be difficult for the social scientists to do in larger, more complex societies.

The characteristics of agent based models are not directly discussed very much in the following chapters because the point of discussing these features here is to briefly introduce why these agent-based models will be a suitable tool to investigate large, complex, social phenomena by social scientists. These models may not provide causal explanations of social phenomena, but they will provide insight into the complex systems of interest.

Another reason for the discussion of these features is to give an example of how these features are applied in a simulation. The point is that these features make these models distinct and useful in aiding the investigations of social phenomena. Showing how these features worked in one case, may give insight on how social scientists can apply agent-based models as a tool in the social sciences.

2.2 Heterogeneity

Heterogeneity is represented in these models when individual (agents) are allowed to differ in "myriad ways—genetically, culturally, by social network by preferences—all of which may change or adapt endogenously over time" (Epstein, 6). Traditional models tend to consider the population as homogenous, meaning that the agents in these models have the same properties as all the other agents. Homogenous agents simply do not represent the individuals in most societies. In most modern societies, individuals are diverse, meaning that individuals differ in their likes and dislikes, their economical status and economical resources, their jobs etc., so in order to represent and simulate that diversity, the feature of heterogeneity would be beneficial to simulations.

Heterogeneity, in the case of the "Artificial Anasazi Project", is represented by Household attributes. The Household attributes and the model "timmng-house hold clocks", allow the food consumption and the union of household to be assigned for each agent. The house hold attributes are assigned as such:

- Five surface rooms or one pit house is considered to represent a single household.
- Each household that is matrilineal and matrilocal consists of 5 individuals.
- Only female marriage and residence location are tracked, although males are included in maize1-consumption calculations.
- Each household *consumes* 160 kg of maize per year per individual.
- Each household can store a maximum of 2 years total corn consumption additional 800 kg can be added to that from the current crop.
- Households use only 64% of the total potential maize yield. (The unutilized production is attributed to fallow, loss to rodents, insects and mildew, and seed for the next planting.)

Model Timing-Household "Clocks"

- Each household has two internal clocks. One clock tracks the number of years a household is in existence and determines when it fissions and dies.
- A household fissions when a daughter marries at the age 16 to form a new household. Birth spacing is at least 2 years.
- A household dies once it reaches it death age, a parameter drawn randomly from a uniform distribution according to model parameter.
- A second clock from April to April reduces the amount of maize in storage by 13.33 kg of maize per month per individual in the household (Epstein, 122)

2.3 Autonomy

Autonomy means "there is no 'top-down', control over individual behavior in agent-based models…as a matter of model specifications, no central controllers or higher authorities are posited *ab initio*"(6). In the case of the Artificial Anasazi, there is no governing body to tell individual households how to organize. Organization and movement rules are triggered by marriages and by agents' subsistence needs. For example, if the amount of stored maize plus the predicted maize production is less than what the agents need to survive, then the agents will move to another farm plot. This characteristic of agent-based models is distinct because most models begin with some mathematical model and use that model to help solve the equations that form part of the underlying model. Agent-based models have no underlying mathematical model, instead there are only rules of behavior that govern the interactions of the agents in the simulation.

2.4 Explicit Space

Epstein produces what he calls an "explicit space." Events usually take place on an explicit space, "which may be a landscape a renewable resource …the main desideratum is that the notion of 'local' be well posed"(6). The landscapes can be completely "imaginary or can capture important aspects of real-world situations"(Epstei90). In agent-based simulations, the explicit space can constantly change which can represent the dynamic changes in environments and customs that are relevant to us as social beings. For example, extreme changes in climate can be simulated in order to investigate how the agents might react which may provide insights into the kinds of reactions people may have in extreme environments.

In the case of the Artificial Anasazi project, the landscape was a reconstruction of the Long House Valley. Digitalizing the explicit space in this case is definitely not imaginary; it is an attempt at a historical recreation of the Long House Valley. There is no yield crop data for the Long House Valley or any nearby or comparable areas. Maize production in the Long House Valley was reconstructed through the integration of information from several sources (Epstein, 94-95). This was necessary in order to speculate the likely production record of maize in the Long House Valley. The sources used include Burns' (1983) and Van West's (1994) dendroclimatic research and the Dolores Archaeological Projects' soil work (Becker and Petersen 1987; Leonhardy and Clay 1985) in southwestern Colorado, E. and T. Karlstrom's (Karlstrom 1983, 1985; Karlstrom and Karlstrom 1986; Karlstrom 1988) soil and geomorphological, Leob's (1991) dendroaricultural research, nearby Black Mesa, Brafield's (1969, 1971) Hopi farming studies, Soil Conservation Services (SCS) soils survey in Apache (Miller and Larsen 1975), and Coconino (Taylor 1983) counties in Arizona.

As a result of not having any data on the soils of the Long House Valley, one or more of the soils in the LHV had to be classified in order to obtain the necessary attributes. Epstein and colleagues had to find potentially analogous soil that was similar to those postulated by the Long House Valley gensoil. (Epstein 95). The potential analogies to the Long House Valley, had to have certain characteristics which included: (1) they had to duplicate the LHV soil families, (2) they had to represent the same elevation range as the floor of LHV, roughly 6,000 to 7,000 feet, (3) they had to have a comparable silt-loam-sand composition and shale derivation and (4) they had to exhibit the default 1"/5" water capacity used in calculating the Tsegi, Palmer Drought Severity indices (PDSI)(Epstein, 96). The Sharps-Pulpit Loam (R7C) and Pulpit Loam (ROHC) were the two soils closest to meeting the criteria. ROHC came closest to LHV gensoil, therefore according to Epstein, the selection of ROHC as the working analog for LHV densoil permitted the use of PDSI to estimate the annual maize crop yield in LHV (96). Epstein suggests that they can estimate the potential annual crop yield for every bit of land, for each of the four environmental zones (96).

2.5 Local Interactions

Local interactions refer to the idea that agents interact with neighbors in the explicit space (Epstein, 6). The actions of one agent will necessarily have an influence on the actions of the other agents in the simulation. . Local interactions are governed by rules of behavior. Agents in these simulations interact with the agents that surround them immediately, which can be seen as an analogous feature of people. Most people interact with their immediate community and rarely have the need to interact with people from distant or outside communities, therefore local interactions seem to be the kind of interactions that most people actually participate in. In the case of the Artificial Anasazi, these are considered to be the rules of behavior for the agents:

- A household fissions when a daughter reaches the age of 16.
- A household moves when the amount of grain is storage in April plus current year's expected yield (based on last year's harvest total) falls below the amount necessary to sustain the household through the coming year.

Identification of agricultural location:

- The location must be currently unfarmed and uninhabited.
- The location must have potential maize sufficient for a minimum harvest of 160 kg per person per year. Future maize production is estimated from that of neighboring sites. If multiple sites satisfy these criteria the location closest to the current residence is selected. If no site meets the criteria the household leaves the valley.

Identification of residential location:

- The residence must be within 1km of the agricultural plot.
- The residential location must be unfarmed (although it may be inhabited, i.e. multi-households sites permitted).

- The residence must be in less productive zone than agricultural land identified in A.

If multiple sites satisfy the above criteria, the location closer to the water resources is selected (Epstein, 122).

2.6 Bounded Rationality

Epstein suggests that there are two components of bounded rationality, which are, bounded information and bounded computing power. Bounded information means that agents do not have access to massive amounts of information or global information, while bounded computing simply means that agents do not have infinite computational power (6).

The importance of these characteristics is that they are characteristics that enable agent-based modeling to represent individual agent actions (7). It seems that this feature is similar to local interactions in the sense that both features are aimed to represent, at least in part, how we as social beings inform ourselves and act on that information.

2.7 Non-Equilibrium Dynamics

Non-equilibrium dynamics is of the utmost importance to agent-based modelers because of their interest in observing large-scale transitions and macroscopic regularities that come out of decentralized local interactions (52). Traditional models in economics for example often reach equilibrium because preferences are assumed not to change. These traditional models are unable to capture the reality of how quickly preferences of agents actually change. Therefore equilibrium is no longer a goal in these simulations and non-equilibrium seems like a much more realistic and appropriate approach. In the case of the Artificial Anasazi project, the large-scale transition that is of interest has to do with the diminishing population of the ancient pueblo people from the Long House Valley and their complete disappearance from that valley.

2.8 Generative Social Science Perspectives

The main purpose of this following section is to illustrate how agent-based models are being used in social science. Some of these perspectives are used to simply justify the use of agent-based models in social science, other perspectives are used to illustrate the possibilities of agent-based models as a tool for social science. Still, other perspectives are used to illustrate some of the possible limitations or challenges of these models. All of these perspectives argue for agent-based simulations being used in the social sciences.

Although there is not more discussion about some of the following topics, these topics are included because it seemed beneficial to get an understanding of some of the assumptions that are being made by generative social scientists. Many of these ideas are addressed in the following chapters and become prime concerns in this thesis; the idea of explanability and interdisciplinary work are examples of the prime concerns. Other ideas are briefly discussed in the following chapters, such as the use of induction is agent-based models and the possibilities of developing and testing of behavior regularities for distinct communities.

2.9 Explainability

Epstein claims that, “agent-based models provide computational demonstrations that a given micro-specification is in fact *sufficient to generate* the macrostructure of interest. Agent-based modelers may use statistics to gauge the generative sufficiency of a given micro-specification, in order to test the agreement between real-world and generated device” (8). For a micro-specification to be a good fit, or a possible *explanandum*, macrostructures must be attainable under repeated application of agent-interaction rules (8). The motto of the generative social science, according to Epstein is, ”if you didn’t grow it, you didn’t explain its emergence” or in other words, “if the distributed interactions of heterogeneous agents can’t generate it (macrostructure), then we haven't explained its emergence"(8,10).

If a micro-specification generates the macrostructure of interest, then it is a candidate explanation, although generating a macrostructure does not necessarily explain its formation (9). If a micro-specification fails to generate the macrostructure of interest, then it is not sufficient and disqualifies it as a possible explanatory candidate. A micro-specification may generate a macrostructure of interest but may do so in an absurd way, which in turn also eliminates the micro-specification as a possible candidate of explanation (Epstein 9, 53). For instance, the Artificial Anasazi may arrive to the actual observed pattern "stumbling around backward and blindfolded" but one would not adopt that picture of individual behavior as explanatory (Epstein, 53). In short, generative sufficiency is necessary but not sufficient condition for explanation. This issue of explainabilty is one reason that I felt that the concept of exploratory value, had to begin to be developed in more detail.

Although agent-based simulations will become an indispensable tool for social scientists who aim to study complex social phenomena, it cannot be assumed that providing the kind of causal explanations of a phenomenon will be as easy as Epstein implies. Plenty of rigorous work is put into the construction of these models and it would be a mistake to assume that these models did not have any scientific value if they fail to give a causal explanation or did not have the "generative sufficiency". Exploratory value, is the possible scientific value of agent based simulations who have failed to give, for whatever reason, a causal explanation.

2.10 Induction in Generative Social Science

Epstein argues that the agent-based models are not inductive in the sense that this term has been classically used in social science. Usually, induction in the social science consists of assembling real-world data and making econometrical relations and approximations, by using techniques of aggregated estimates (Epstein,10, 66). In generative social science, a model can be tested a large numbers of times, each time adding and subtracting micro-specifications, and then one can "collect clean data and build up a robust statistical portrait of model output"(Epstein,66). The goal here is twofold according to Epstein, first, "to understand one's model when closed form analytical expression are not in hand;"and second, "to explain observed statistical regularities" (66). While the large collections of model realizations are "*each realization is a strict deduction*", "it is *induction over a sample distribution of theorems*, in fact" (66). The use of induction becomes an issue in section 4.1.6 in the response to Grüne-Yanoff's claim, that the ethnographies that Epstein used were too vague.

2.11 Deduction in Generative Social Science.

The generativist, Epstein argues, accounts for an "observation as explained precisely when we can *deduce the proposition expressing that observation from other more general propositions"*(10). Epstein argues that propositions are the only objects that we can deduce. In turn, "we explain an empirical regularity when that regularity is rendered as a proposition and that proposition is deduced from premises we accept" (56). From Newton's Second Law and the Law of Universal Gravitation, one can deduce that "the acceleration of a freely falling body near the surface is independent of its mass" (10-11).

According to Epstein, generativists are requiring that all propositions expressed be observed and generated in an agent based computational model (11). Because all programs can be expressed in "recursive functions and are mechanically (effectively) computable-in principle by hand with pencil and paper," Epstein argues, that they "can legitimately claim that they [computational models] are strictly deducible"(56). For every state generated in an agent-based model, one can deduce a theorem, for every pattern observed on agent-based model, there is a theorem to explain that pattern (56). Epstein states that for "every computation, there is a corresponding logical deduction", technically "*generative implies deductive*". But the opposite is not true, deductive does not imply generative (11-12).

2.12 Interdisciplinary Work

In reality social processes and social systems are not easily broken down into categories. Academic disciplines often share systems of interest so it would seem prudent to have interdisciplinary work into these systems of interest. Epstein suggests that agent-based models offer new techniques for interdisciplinary studies. Epstein believes that many different ideas or concepts can be studied at once because different micro-specifications can be active at different times or they can all be active at once, "when an initial population of such agents is released into an artificial environment in which, and with which, they interact, the resulting artificial society unavoidably links demography, economics, cultural adaptation, genetic evolution, combat, environmental effects and epidemiology. *Because the individual is multi-dimensional, so is the society* (18). But agent-based modelers can also relax certain micro-specifications in order to test robustness of theories. For example, in microeconomic theory, it is assumed that individual preferences are fixed for a lifetime. But what if they are not fixed, what if the preferences vary culturally, Epstein explains that this change of assumption in microeconomic theory, leads to "far-from-equilibrium markets"(20). This becomes a major issue at the end of chapter four where it is argued that interdisciplinary work will be necessary in constructing, testing, adjusting, and analyzing the results of agent-based simulations.

2.13 Behavioral Sciences

Epstein suggests that the agent-based model will help behavioral sciences, "project up" from individual behavioral to the macro level. Epstein states that "even perfect knowledge of individual decisions rules does not allow us to predict macroscopic structures, therefore the behavioral sciences will be able to implement agent-based models which will help in "projecting up" to the macroscopic structures (21).

Agent-based models may also provide the behavioral scientists research with unanticipated assumptions regarding individual behavior, then it will be "fruitful to design laboratory experiments to test hypotheses arising from the unexpected generative sufficiency of certain rules" (21). Epstein suggests that, "some, apparently bizarre, system of individual agents rules may generate macrostructures that mimic the observed ones. Is it possible that those are, in fact, the operative micro-rules?" (21).

Another way in which the agent-based models can help the behavioral sciences is, that if there are two micro-specifications that are completing hypotheses of individual behavior, then experiments. Using agent-based models can be designed in order to identify the better hypothesis (20). These unanticipated assumptions, described by Epstein above, are partially what is meant when I argue that agent-based simulations might give us lines of investigations that otherwise would not have been investigated. This is also what I meant when I argue that these simulations can provide possibilities of investigation where there are few otherwise available options.

2.14 Decoupling

The ability to decouple individual rationality from macroscopic equilibrium is one of the features of agent-based models. Accordingly, rationality on behalf of agents will affect the rate at which any equilibrium appears. The issue according to Epstein is, "*how little is enough to generate the macro equilibrium*" (22). Epstein states that the "manipulation of uncritical imitative impulses maybe more effective getting to a desired macro equilibrium than policies based on individual rationality" (22). Epstein considers that agent-based models, as policy application, may suggest ways to operate on-or ’tip’-ethnic animosity itself” (23). Epstein believes that agent-based models can aid in social contagion studies, where the affects of “fad creation” can be studied.

2.15 Foundational Issues

Agent-based models, according to Epstein, allow us to study long-lived transient behavior. However, “depending on the number of agents and their memory lengths-the waiting time to transit from an inequitable regime to the equitable one may be astronomically long” (24).The number of agents and the agent’s memory length are factors in the transit of equilibrium so then, the time it takes to transit becomes an issue of interest. Epstein suggests that like “satisfiability, or truth table validity in sentential logic, these problems are in principle decidable (that is, the equilibria are effectively computable), but not on time scales of interest to humans” (25).

2.16 Gödelian Limits

Kurt Gödel proved that *all sufficiently rich mathematical systems are incomplete;* in essence there are true statements that cannot be determined false within a sufficiently rich system such as proved by Russell in Principia Mathematica. Epstein will argue that there are problems in the social sciences that are undecidable in principle. Rabin (1957) states that “intuitively, our results means that there are games in which the player who in theory can always win, cannot do so in practice because it is impossible to supply him with effective instructions regarding how he should play in order to win”(26). Epstein is suggesting, that social configurations are not complete systems, many social sciences have implicit assumptions, therefore there will be social configurations that are undecidable in principle as well (61-62).

2.17 Rational Choice Theory

Epstein believes that one of the limitations with Rational Choice Theory is the belief that irrational agents, who have no central control, will produce "suboptimality at the aggregate level… the invisible hand requires rational fingers" (26). Epstein contends that Rational Choice Theory assumes an idealization of individual rationality, a sort *of homo economicus*. The problem with an idealization like *homo economicus* is that in actuality, it behaves very differently from *homo sapiens,* therefore in social science, "it is appropriate to ask whether the idealization of individual rationality in fact illuminates more than it obscures"(62-63). For Epstein, it is equilibrium attained from decentralized local interactions of heterogeneous boundedly rational actors that is much more interesting and insightful of macro social phenomena (27).

2.18 Equations Versus Agent-Based Models

Epstein argues that for every computational model, there is an equivalent equation (27). This is one of the main contributions of computer science, according to Humphreys. Humphreys argues that one of the reasons that we must reevaluate the epistemology of philosophy of science is because computers give us solutions to equations that otherwise would be intractable.

According to the Church-Turing Thesis, a Turing machine can execute any computation and in turn for every Turing machine there is a "unique corresponding and equivalent Partial Recursive Functions"(27). In principle, for all computational models there exists an equivalent equation (involving recursive functions) (27).

Epstein argues that a researcher's choice of agents versus equations always hinges on the objective of his or her analysis.

2.19 Challenges

Epstein believes that one of the challenges in the field of computational models is that there are no standards for model comparison or replications of results (29). Updating of agents, for example, will have radically different results, when the updates are synchronous versus when the updates are asynchronous. Another important issue arises when one tries to identify which postulated rules are credited for the observed macro-phenomena. Epstein explains that “we have no efficient method of searching the space of possible individual rules for those that exhibit generative power” (30).

Epstein also suggests that another problem agent-based modesl has to do with encoding the “vast space of possible individual rules (not to mention the raw computational challenge of searching it once encoded)” (30). I will argue that the vast possibilities of rules of behavior can be narrowed with the use of ethnographies, statistics, and other methods developed by and shared in the social sciences.

Assessing sensitivity in agent-based models is also a challenge according to Epstein. In Dynamics Theory, sensitivity refers to small changes in initial inputs, having large effects on the outputs; meaning small changes in micro-specifications will have large effects on the resulting macrostructures. Epstein suggests “metricizing” agent-based models. He argues that without some metric, “we cannot develop the analogue, for agent-based models, of structural stability-or equivalently”(31). Other challenges include building community and sharing results, foundational, procedural, and terminological problems (31).

2.20 Generative Social Science: Running the Simulation

The next two sections are relevant because I believe that philosophers can only make commentary on conceptual aspects of agent-based simulation, if they have also concerned themselves with matters of application. One of the main ideas of this thesis is to argue that the scientific value of agent-based models or "exploratory value," will consist of: the construction, testing, adjustments, and analysis of results. This in turn involves investigation into the application process. I will argue that no causal explanations can be derived from the use of this simulation in isolation; instead exploratory value of the simulation is necessary.

One could argue that the following description of the running process, and analysis of the simulation results, could be enough for philosophers to continue with their traditional concerns. There are important insights that come from the actual use of agent-based simulation that would not be attained through descriptions alone. Nonetheless, the Artificial Anasazi Project is a good example of the move from purely conceptual models to agent-based models.

The Artificial Anasazi Project aimed to recreate the actual spatio-temporal history of the Long House Valley in Northeastern Arizona between roughly 1800 B.C. and A.D. 1300. The Long House Valley is a topographically discrete 96 km2. The recreation was constructed from research of the Long House Valley Project, a research effort of the Museum of Northern Arizona and the Laboratory of Tree-Ring Research at the University of Arizona, which involved a 100 percent survey of the valley (Epstein, 91-92). The data was extracted from computer files of the Southwestern Anthropological Research Group, SARG. SARG was "an effort at a large-scale data accumulation and management and cooperative research" (Epstein 93). The data was downloaded from SARG master file, "modified through the elimination of many categories of data deemed extraneous" and then imported into Artificial Anasazi software (Epstein, 93). The data on the location and site are used as "the referents against which the simulations are evaluated" (Epstein, 93).

The simulations take place on a landscape, which varies annually according to maize production potential (Epstein, 93). The variations are based on "empirical reconstructions of low and high frequency paleoenvironmental variability in the area. The production values represent as closely as possible the actual production potential of various segments of the Long House Valley environment over the last 1,600 years" (Epstein, 93).

The Long House Valley area is well suited for agent-based modeling because of four reasons according to Epstein. The first reason is that the area is topographically bounded, self-contained" landscape that can be simulated by a computer (93). Second, a large paleoenvironmental record based on alluvial geomorphology, palynology and dendroclimatology allows for a reconstruction of the annual fluctuations of potential maize production (93). Third, ethnographies of people in the region allow for possible rules of behavior for the agents. Fourth, archaeological research involving 100 percent survey and limited excavations create a database of real-world behavior of the last 2,000 years, which in turn, becomes the target that models aim at. The introduction of maize brought forward a food-producing economy and Anasazi culture, which continued until the abandonment of the area about A.D. 1300. (Epstein 93-94). Reconstruction of environmental variability, is based on relationships of measurements, such as "the rise and fall of alluvial groundwater and the deposition and erosion of floodplain sediments", the measurements were used to create a "dynamic landscape of annual maize production, in kilograms, for each hectare in the study area for the period A.D. 382 to 1400" (Epstein 94).

The study period runs from A.D. 800 to 1350. The historically known number of agents was used but not the historical settlement locations (Epstein 101). Each household carries out its full range of behavior i.e. moving consuming, reproducing, storing food and if needed, leaving, each year. If the agents make fortunate choices, the household produces enough food and the household survives another year, if not, the household runs out of food and is either removed from the simulation, as in the cases of death of emigration (Epstein, 101). Many runs of the simulation, involving altered initial conditions, parameters and random number generators should be pre-formed in order to test the models robustness, according to Epstein (101). Epstein states that some of the models outputs can be characterized statistically and compared to LHV data, outputs that cannot be easily characterized "can be visually compared to real-world patterns" (101)

2.21 Simulation Results Compared with Archaeological Data

A simulated map of household residence and field locations is run alongside a map of the corresponding archaeological and environmental data in order to ease assessment between historical and simulated population dynamics and residence location (Epstein, 102).

According to Epstein, both maps typically exhibit the similar relative variations, meaning that both maps show an increase in population until about A.D. 900 then leveling off in the tenth century, then hitting a major population increase between 1000 and 1050 A.D., from 1050 to 1150 the population levels off again (102). In the middle of 1100s there is a drop in population until a resurgence in the late 1100's, which hits a peak in the thirteenth century and finally a major crash in the population in the late 1200s(102). The maps also demonstrate important qualitative differences such as "a greater and more prolonged simulated population decline in the twelfth century, a more immediate, more gradual and relatively higher post-1150 recovery in the archaeological population, a slightly earlier thirteenth century decline in the simulated curve, and the failure of the simulated curve to drop to zero at 1300"(102). Epstein argues that on the large scale, the simulations replicate important aspects of the settlement history of the Long House Valley with uncanny accuracy. The patterns of aggregation and dispersal almost duplicate the settlement history of the eastern Kayenta Anasazi area.

The settlement patterns and shifts of the Ancient Pueblo people is due to low-and high-frequency environmental fluctuations and according to Epstein, it is "clear that the Artificial Anasazi Project successfully captured the dynamic relationship between settlement aggregation-dispersal and low- and high-frequency environmental variability in the study area" (106).

According to Epstein, in general, there is a qualitative agreement between the simulated map and historically corresponding map. However, there are significant quantitative differences in household numbers and settlement sizes. Epstein states that the total population and the individual settlement sizes are much larger in simulation than what archaeologists assume to have been the actual case, populations gathering together occur earlier and more frequently in the simulations than in the historical record (102). The simulations gather together more households into a single residential location than what is historically believed. The actual Ancient Pueblo people distributed "members of residential unit across a number of discrete but spatially clustered habitation sites" (105).

A combination of low-and high-frequency environmental degradation, rapid arroyo cutting, low and high frequency environmental degradation, depleting alluvial water tables, the Great Drought of 1276-1299 and a breakdown "spatial coherence o seasonal precipitation, created the "most severe, subsistence crisis of the nearly 2000 years of paleoenvironmental record" (106). In the simulations, the agents never completely abandon the Long House Valley but Epstein believes that "the behavior of the artificial aggregation after 1250 is extremely instructive about the possibilities of human occupancy of the area during intervals of environmental deterioration and the high population densities (Epstein, 106). In the simulations, the number of larger settlements did drop abruptly but they did not disappear completely, whereas small to medium size settlements did not change, even when environmental factors became the most stressful and in fact, increased a great deal after 1300 (Epstein, 106). Epstein believes that these results, along with paleoenvironmental data, are evidence that the Long House Valley could have supported a reduced population. This implies that the Ancient Pueblo people living in the region could have remained in the region had they broken up into smaller communities (106).

Epstein is arguing that the simulations provide insight to the reasons why the Ancient Pueblo people may have disappeared from the Long House Valley. He suggests that the failure of the simulation to replicate the correct number of settlements is an indication that environmental factors only account for a reduced population and not the complete abandonment of the Valley. Epstein is going to suggest that environmental stress may have been partially responsible for the Ancient Pueblo's people departure but cultural factors must have also played role. He believes that "The delicate balance between environmental 'push' factors and non-environmental (cultural) 'pull' factors suggested by the artificial Long House Valley results is compatible with long-standing, archaeological untestable hypotheses about the real Anasazi world" (107).

Epstein briefly discusses side-by-side photographs of three particular years. The photographs are of what he refers to as the "real history", meaning a reconstruction of what archaeological evidence shows to be the true history of the Ancient Pueblo people's, and the simulations, where the agents have been given rules of behavior. The comparison of the three years seems to be a kind of synopsis of the overall performance of the simulations. The first year that Epstein compares is the year 1000, he comments that while the number of simulated settlements are too many, far more than what is actually believed to have existed. The simulation did correctly reflect where the real sites where located, namely, "along the periphery of the flood plain throughout the entire valley"(107). Apart from a few similarities, Epstein suggests that the simulations only performed moderately well for that particular period (107).

The next year that Epstein compares is 1444. Epstein suggests that the simulation mimics the spread of settlements throughout the valley, including the clustering of settlements into five groups. He also suggests that the simulation mimics the scattered sites in the non-agricultural uplands and also accurately simulates the settlements in the appropriate environmental zones. Although, the simulation correctly portrayed many of the locations of settlements, the major difference between the simulations and the 'real' situation is the greater amounts and size of settlements in the simulation, in particular in the north-central uplands and upper Kin Biko (107,109).

The last year that Epstein discusses is 1261, where he suggests that the simulation "spectacularly duplicates the abandonment of the southern half of the valley as a place of residence and the concentration of the population along the northwestern edge of the flood plain near the remaining patches of productive farmland" (109). Although the simulation places many of the correct number of sites in the correct parts of the valley, significant discrepancies are found in the Midvalley Floor where there was an absence of sites, as well as the simulation placing too many settlements in the Kin Biko. (109-110). Epstein also believes that the simulation, in accordance with some empirical evidence (Dean et al. 1978), correctly illustrated the order of which different environmental zones were exploited by the real Anasazi (110). Epstein states that although the simulated Anasazi do not completely abandon the Long House Valley, the simulation illustrates a major population decline in the same period that as the real Ancient Pueblo people's exodus. The simulated Anasazi that did not abandon the Long House

Valley after 1300, continued to locate fields in the Long House Valley under "vastly altered environmental conditions" (112).

Epstein believes that these simulations show that the Ancient Pueblo people did not have to totally abandon the Long House Valley due to environmental deterioration. He goes on to state that a substantial fraction of the population could have stayed behind if they would have dispersed into small settlements and moved to land still suitable for agriculture. Because the population "chose" not to stay behind, Epstein believes that cultural "pull" factors "were drawing them away from their homeland" (112-113).

Other conclusions that Epstein reaches are that the simulations support the "predictions" that under different environmental conditions, different kinds of farmlands are used. Epstein also concludes that the simulation "illuminates" the importance "interactions among various demographic and environmental factors in the processes of socio-cultural stability, variation and change"(113).

Chapter 3: Literature on Computational Science, Simulations and Computer Models

In this chapter I introduce my view of the distinctive value of computational science. I will explain what I mean by the exploratory value of computational models and will try to make the point that this value should be taken seriously by social scientists.

I provide a preliminary definition of what I refer to as, the exploratory value of simulations. Let us make an initial assumption: exploratory value is not the same as a causal explanation, meaning that the exploratory value of a simulation is not the claim that the complete and particular causes of phenomena have been determined. In order for us to begin a discussion on simulations, it is necessary to familiarize ourselves with some of the concerns with simulations such as, under what circumstances are simulations used and how can we judge the reliability of simulations. The goal in this section will be to introduce some of the answers to these questions.

3.1 Working Definition of Exploratory Value

According to Winsberg, when data for a system of interest is sparse, a large range of sources is needed to build and analyze the results of simulation. He concludes that the results from the simulation should replace experiments and observation as sources of data about the world, (Winsberg, 121). But if we replace experimental and observational data, with the data that we gained from the simulations, then the next intuitive question may be, how do simulations gain credibility or how do we come to trust the data given to us by a simulation?

Part of the answer of how a simulation gets credibility and how the results of a simulation are sanctioned is that simulations must use large range of sources. The large range of resources include according to Winsberg, the mechanical models, calculation techniques, background knowledge of “principles for model building” and the skills with which the simulation was implemented. (30-70).

Winsberg suggests that another way in which simulations gain credibility is through a maturing process. Winsberg, who uses the same line of reasoning as Ian Hacking (1992), argues that simulations like experiments mature through a process, which involves testing, adjustments and error correction (44). I agree with Winsberg when he suggests that there are two separate claims that can be made when we adopt Hacking’s line of argumentation.

The first claim is that simulations are analogous to experiments because they change over time (45). As time passes, experiments and instruments used in the gathering empirical evidence, are adjusted for more precise information and to compensate for new conditions. The same can be said about simulations; they too evolve and mature over an extended period of time and use. Because a simulation is going to need to be adjusted, it implies that the simulation has failed to give a complete causal account of the system of interest. Meaning that the simulation did not produce the phenomena of interest but did produce results that inspired new ideas on how to adjust the simulation. The adjustments could be, for example, made on the data implemented in the simulation, or the adjustments could be made to the construction and formal properties of the simulation itself. The point is that the failure of the simulation has led to new insights and has opened a new line of investigation to the researchers involved.

Even if the simulation fails, it should not be disregarded completely. This is a feature of a simulation of having exploratory value. For example, the Artificial Anasazi simulation has exploratory value because Epstein and his colleagues concluded that more than likely the departure of the Ancient Pueblo people was partially caused by some sort of cultural rule that was not implemented in the simulation. Till Grüne-Yanoff refers to this as the “functional component", a point of discussion later in this chapter. So adjustments of simulations are characteristics of a maturing process and a feature of exploratory value.

Another implication of these adjustments is that they will take time to be implemented and retested. As Winsberg claims, the evolving of the simulation will occur over an extended period of time, meaning that the exploratory value of a simulation can provide more than one kind of insight and can give different insights at different times.

The second claim is that simulations comparable to experiments are "self-vindicating." Over time, simulations like experiments have their own credentials and these credentials are a result of successful techniques and applications of the simulations (Winsberg 45).

The Artificial Anasazi simulations is an example of how a simulation can be successful by fitting well with, at least some of our previous knowledge. In the Artificial Anasazi simulation, the agents displayed patterns of migration and dispersal that closely resembled the actual patterns of the Ancient Pueblo people. A simulation can investigate rules of behavior that will allow for large macro-phenomena to emerge. This is one of the fundamental ideas behind the Generative Social Science approach and behind fields of research, such as Complexity. Each time that the results from a simulation are successful, their credibility as a reliable source of data, increases as well (Winsberg, 70). Although it may have failed at producing the phenomena of interest, namely the disappearance of agents from the landscape, it fit well with at least some of our previous data.

The idea that simulations have their own credentials and become more reliable with applications can be demonstrated by using the Artificial Anasazi simulation as an example. If someday one wanted to learn more about the Ancient Pueblo people's patterns of migration and dispersal, then one could start with the Epstein's simulation. Then one would continue by making the proper adjustments to the simulation in order to accommodate for one's own study. The point here is that one would not have to start from new. A simulation that is relatively successful at replicating some of the migration and dispersal patterns of Ancient Pueblo people is in principle, available to begin adjusting and experimenting with. The simulation demonstrates characteristics of an exploratory value through the maturing process described above. If a simulation needs adjustments over time and if it is self-vindicating, meaning that the adjustments to the simulations come from insights resulting from the simulation itself, then the exploratory value is evident.

Winsberg argues that data from simulations should replace experiments and observation as sources of data about the world because the simulation has incorporated information from a wide range of sources. Simulationists, according to Winsberg, try to maximize fidelity to theory, to mathematical rigor, to physical intuition, known empirical results, mechanical models, and calculation techniques that have proven to be reliable (45) In other words, the data from other experiments and observations were included in the simulation, and the simulations have produced results that give more details about the system of interest. Therefore the data produced by the simulation replaces all other previous data, according to Winsberg.

Now that we have an understanding of the maturing process that is necessary for simulations to gain creditability and why data from simulations should replace data from experiments and observations, we should recall some of the conclusions that Epstein draws from the Artificial Anasazi simulation and see how exploratory value plays a role in that simulation.

Epstein contends that the Artificial Anasazi simulation gives clear evidence that the environmental deterioration alone could not have been the reason for the complete abandonment of the Long House Valley. Epstein asserts that environmental stress was only partially to blame for the abandonment. In addition, he suggests that the simulation shows that not all the Ancient Pueblo people had to abandon the valley and that a substantial fraction of the population could have stayed in the valley, by dividing into small settlements and dispersing into land that was still suitable for agriculture (113-112). Epstein argues that the simulation clearly indicates that environmental factors alone where not to blame for the disappearance of the Ancient Pueblo people, which in turn implies that cultural factors must have played a role in the exodus as well. Epstein suggests that the environmental stress was a kind of "push" factor, while socio-cultural factors were the "pull" factors, which lead to the abandonment of the valley. This last point is made evident when Epstein states that "environmental degradation was not responsible for the complete abandonment of the valley and that other, undoubtedly social, factors were involved in the final emigration" (141).

If we recall, one form of gaining credentials for a simulation, is through results from successful techniques and applications of that simulation. The need for credentials is going to be necessary when trying to support the causal explanations of the phenomena. However, if we are observing the simulation for its exploratory value, then it is clear that few credentials are needed. What we are suggesting is that the simulation has enough redeeming qualities that allow for a continual use of that simulation, illustrating the self-vindicating potential, which implies the exploratory value of the simulation. While Epstein will claim that the simulation closely mimics the aggregation and dispersal pattern that was initially mapped by archaeological evidence, in reality the aggregation and dispersal pattern of interest, the target system, the macro-phenomena which they hoped to "explain," namely, the disappearance of the Ancient Pueblo people from the Long House Valley, was never mimicked by the simulation.

One point that Epstein and I disagree on is the *explanatory value* of the Artificial Anasazi simulation. I do not believe that this particular simulation has the explanatory value that Epstein suggests. The simulation never mimicked the phenomena that was trying to be explained, therefore the rules of behavior that where given to the agents in the simulation were also not explained. In fact we could argue that the aggregation and settlement patterns of the Ancient Pueblo people had already been known to us through archaeological evidence, hydroclimatelogy and paleontology evidence. The simulation was being compared to a digital map that reflected the "true history" of the Ancient Pueblo people. So we already knew the aggregation and settlement patterns of the Ancient Pueblo people and Epstein used that data to compare his model with, so quite literally the settlement patterns were already mapped. This last point is clear when we take into account that "archaeology is the only social science that has access to data of sufficient time depths to reveal long-term changes in patterned human behavior" (Gilbert & Conte, 1995). The simulation missed the target system which in turn suggests that the rules of behavior given to the agents in the simulation where not completely explained and the aggregation and dispersal patterns that resulted from the simulation were already know to us by more mature, trustworthy and reliable sources, namely archaeological evidence, the same evidence that Epstein used in the testing and building of his model. If Winsberg is correct and simulations are supposed to replace the existing data, then we can imply that the simulation failed at that as well. We can suggest that the simulation in all actuality did not provide us with any new information about the aggregation and dispersal patterns of the Ancient Pueblo people.

Epstein's "principles of modeling" are not in question, and as mention earlier, Epstein's knowledge of the target system is not being questioned either. I believe that Epstein and his colleagues made exhaustive efforts to research and replicate the environmental data given to them by previous archaeological evidence. What is questionable is whether Epstein *explained* the target system and the answer is no, he did not. So although Epstein and his colleagues may have had all the adequate data about the target system and although the "principles of modeling" may have also been applied intact, the inability for the simulation to fit in well with *all* of our previous data and intuitions that prevents the simulation being explanatory. It should be clear that gaining credentials is not just a matter of the simulations fitting well with our prior intuitions and data. It is the cohesion of prior successes matched with the proper background knowledge of the target system and the proper knowledge of the "principles of modeling" that ultimately gives us confidence in the model. As Winsberg stated, "it is the simultaneous confluence of these efforts, rather than the establishment of each one separately that ultimately gives us confidence in the result" (23).

By questioning the credibility of the Artificial Anasazi simulation, my intention is to bring to question the maturity of the model. The model does not have the efficient amount of time and use to evolve and to recalibrated the simulation; the process that a simulation must go through in order to gain credibility has not occurred in the case of the Artificial Anasazi project. The simulations must be retested and retooled numerous times, by other numerous modelers, before the data given by the simulation can be trusted. It seems unrealistic to believe that a long lasting archeological debate can be settled by a novice simulation[2]. The simulation must have a degree of prior adjustments made through an extended period of time before it can be relied upon as being trustworthy. Although the trustworthiness of the simulation is in question because of the of amount of time and adjustments that have been put into it, the exploratory value is intact because we understand that the simulation cannot be trusted to give a full causal account.

[2] The debate that I am referring to has to do with whether or not the disappearance of the Ancient Pueblo people was caused solely by environmental degradation. One advocate of the idea that the Ancient Pueblo people disappeared because of environmental degradation is Jared Diamond and he makes the argument in his book entitled *Collapse*

However the simulation will have insight into possible new lines of investigation, such as, cultural “pull” factors that may have been involved in the exodus of the Ancient Pueblos people.

An overview of these ideas suggest that the exploratory value of a simulation can be defined when a simulation has failed at giving a causal account but has given new insight into the system of interest. The exploratory value of the simulation can also be defined as the deficiencies in construction of the simulation, implying that exploratory value can also be assigned when the possibilities of further research has resulted from the simulation. The simulation must have been adjusted over time and the data produced by the simulation does not have to replace all of our previous data.

3.2 Humphreys’s Idea of “Extending Ourselves” as Part of Exploratory Value

Another reason an agent-based model may have exploratory value, has to do with the kind of scientific investigation that may continue because of agent-based simulations. This idea is reflective of Humphreys' idea of “extending ourselves”. It seems clear that one of the major reasons that researchers employ simulations is because the data for the target system is sparse or because it would be extremely difficult to experiment on or observe the system under investigation. Due to sparseness in data and because of the difficulty in direct experimentation and observation of the system of interest, many projects in science could possibly end up at a standstill and researchers would be unable to continue with their work. Agent-based simulations can and have assisted scientists in furthering research that otherwise would have been physically impossible. If a simulation can assist in furthering research by giving new insights or by aiding experimentation, by leading us to reevaluate assumptions, and by sparking creative approaches, then simulations employed in any of these ways will also have exploratory value.

This is what Jesse L. Voss had in mind when he attempted to reconstruct the Artificial Anasazi simulation and add a protocol for mate selection. Although Voss admits that the Artificial Anasazi simulations are limited in their predictive power, he believes that simulation could "be used to create a series of alternatives scenarios, etc. These would hopefully be used by different disciplines to answer different questions and open up whole new possibilities for inquiry and research utilization"(15). Voss is describing the exploratory value of the Artificial Anasazi simulation, when he stats that small contributions can lead to new possibilities for inquiry and research utilization. If we were to replace Epstein's assertion that the Artificial Anasazi simulation had explanatory value and suggested a more humble conclusion such as, the simulation may have exploratory value, then it would seem that the model was not a complete failure and it does have some scientific value.

Although the Anasazi simulation did not give a causal account of the system of interest, it would be a mistake to completely disregard the simulation as some sort of toy with no scientific value because it can potentially give us insight into phenomena of interest. Humphreys makes this idea clear when he states that agent–based simulations, "instead of giving a realistic account of some social phenomena, gives us *insight* into the social phenomena modeled" (2004,132). If the last quote is correct, then we can see that the Artificial Anasazi model does have scientific value. If we reevaluate what simulations can do for our sciences then we can see how simulations will be important tools for investigating complex, dynamical systems by giving scientists counter intuitive possibilities of explanation. Agent-based models will be especially useful in exploring micro-actions that result in macro-phenomena because these kinds of simulations, like Adjustable-parameter models "allow the parameters and functional forms to be adjusted to fit data" (Humphreys 2004, 132) therefore different rules of behavior can be explored and tested.

Exploratory value, in the case of agent-based models, also implies that there is further work to be done and that the rules of behavior for agents within the model may be incomplete or incorrect. On the other hand, explanatory value implies that the rules of behavior for a phenomena of interest have been "explained", which in turn implies a complete understanding of the rules of behavior that cause the macro-phenomena to emerge. No need for further investigation, we know what behaviors caused macro-phenomena in question. Obviously there is a need for more investigation when it comes to the Artificial Anasazi simulation because as we clearly know by now, the main dispersal pattern of interest was not simulated and therefore further investigation into other rules of behavior must be described and tested. In other terms, exploratory implies need for improvement and not completely understood as was the case for the Artificial Anasazi simulation.

Another reason why the Artificial Anasazi simulation may retain some scientific value has to do with Humphreys's idea of epistemology shifting from human based to becoming technologically based. Humphreys argues that technological achievements enhance our "native cognitive abilities" and are now a routine part of our sciences (2004, 8). Humphreys believes that there are three particular ways in which we extend our native cognitive abilities in science, which are extrapolation, argumentation and conversion. For the sake of this discussion, I will only need to review two: argumentation and conversion.

Argumentation is another form of extending ourselves or enhancing our natural abilities, according to Humphreys. Argumentation occurs when we attempt to inquire into phenomena that we are naturally not able to detect or account for (2004, 4). Computational devices allow us to go beyond our natural mathematical talents, increasing the speed in which we can perform mathematical endeavors (ibid). Computational devices and observations, supplement one another when those devices are used to give us improved access to the natural world (ibid).

Conversion, according to Humphreys, is another way in which we enhance our natural abilities. Conversion occurs when an instrument changes our access to a phenomenon from one sensory modality to another (ibid).

When we attempt to "extend ourselves," we are attempting to enhance our natural epistemic capabilities in order to investigate and research phenomena that may otherwise be infeasible, such as the inner convective flow of a star. We cannot physically probe the star, therefore a simulation must be constructed. Rules of behavior, especially in historical cases, seems be another kind of subject that would be beyond our natural capabilities of investigating. One of the problems of having sparse data, according to Humphreys, is that results from simulations can never be denied or confirmed based solely on observations. However, in order for simulation to have exploratory value, the results do not have to be confirmed or denied. In fact, like we have argued, the results can be false and still give insight and consequently exploratory value to a simulation. Humphreys supports the idea that a simulation can be false but still capture reality, when he states that, "overall, the model is almost always false…yet despite this overall falsity, parts of the model often are true and some of the terms do refer"(2004,82).

Epstein through the use of Argumentation and Conversion was able to theoretically explore rules of behavior for people who disappeared long ago. Even though Epstein was not able to identify the exact rules of behavior that caused the disappearance, he was able to *extend* *his natural cognitive abilities* in order explore possible rules of behavior that may have been responsible for some of the movements and settlements of the Ancient Pueblo people.

The ability to explore and inquire is a benchmark of the scientific method and simulations give us the opportunity to begin exploring and inquiring into systems that we either had no access to or whose complex structure was much too complicated for our unaided abilities to comprehend or even begin to explore. The value of the Artificial Anasazi simulation is evident; Humphreys supports the idea that simulations can be exploratory tools because the simulation provides us with the possibilities of continued research. Continued research that utilizes the Artificial Anasazi simulation may give us a better understanding of possible reasons and possible rules of behavior that may have contributed to the sudden disappearance of the ancient Pueblo people of the Long House Valley.

The point here is that by extending ourselves or enhancing our natural capabilities by employing simulations like the Ancient Anasazi simulation, we are afforded a method of investigation and exploration. Humphreys would also support the idea that the Artificial Anasazi does have exploratory value in the sense that it has given us insights into some of the possible rules of behavior that may have been responsible for the settlement and dispersal trends known of the Ancient Pueblo people. The simulation has also given us another method to explore social phenomena of a group of people who have long ceased to exist. Finally the simulation has initiated investigation into possible rules of behavior of the Ancient Pueblo people. From the results given by the simulation, other possible rules of behavior can be investigated and tested. So although the rules of behavior in this particular simulation where incorrect, further investigation into the correct rules of behavior has been aided through the process of elimination.

After reviewing Humphreys and because we are describing exploratory value and not causal explanations we made some adjustments to a few of Winsberg's ideas in order to be consistent with previous ideas of exploratory value. One adjustment is that, for a simulation to have exploratory value, there is no need for the results of the simulations to replace all the pervious data. We understand that some of the data within the simulations may be erroneous, because it would difficult to know exactly what data is causing the error. It would be a mistake to simply replace a previous data with the new data from the simulation. Another adjustment is that because we are searching for exploratory value and not causal value then we will not have confidence in all the parts of the simulation to produce credible results. Again, it seems that when Winsberg is talking about credible results, he is describing credible *causal* results. Because we are not interested in causal results and because we understand that our simulations have failed in part, we then understand that adjustments will have to be made to the simulations in someway, therefore we will not have complete confidence in all the parts of the simulations.

Every time that we use technology to probe into a system of interest that we would not be able to probe into without the use of that technology, we are extending ourselves. This seems to be a subtle but significant addition to our concept of exploratory value, specifically computer based technology.

3.3 Grüne-Yanoff and Potential Functional Analysis

Now I will look at Till Grüne-Yanoff's ideas on simulations and see how his ideas on "potential functional analysis" fit in with our idealization of exploratory value.

Grüne-Yanoff maintains that there are at least three differences between potential *causal explanations* and potential *functional analysis* and a clear understanding of these differences will help explain the difference between casual explanations n and exploratory value. I believe that my claim that agent-based simulations can have exploratory value is relatively similar to Grüne-Yanoff's idea of potential functional analysis. Therefore, I will briefly describe Grüne-Yanoff's idea of potential functional analysis and see if any adjustments are needed in order to apply it towards a better description of exploratory value.

The first important difference between causal explanations and potential functional analysis, according to Grüne-Yanoff is that "functional analysis individuates not according to possible factors or mechanisms, but according to possible functions" (551). According to Grüne-Yanoff, simulations can pinpoint the "functional component" missing in an existing model, which in turn specifies the role of the missing element in the context of the existing model. In order for a simulation to have potential functional analysis, the simulation must be able to give us insight into the "functional components" that may be missing in the simulation.

The second important difference between causal explanations and potential functional analysis, according to Grüne-Yanoff is that, possible functional analyses are transferable across different causal contexts. The third important difference is that, functional analysis shows how lower-level capacities constitute higher-level capacities (552). I believe that these ideas of potential functional analysis could be applied to the concept of exploratory value but as in the case of Winsberg's account, adjustments may be needed in order to be consistent with a working concept of exploratory value.

I will argue that the Artificial Anasazi simulation displayed the three ideas found in a potential functional analysis but I believe I had to evolve and adjust the current definition of exploratory value. In order to see what adjustments will be made let us examine Grüne-Yanoff's assertions on simulations, specifically the Artificial Anasazi simulation.

Grüne-Yanoff does not believe that the Artificial Anasazi simulation has any explanatory value. Grüne-Yanoff objects to the claim that agent-based simulations can help explore for possible rules of behavior. In particular, he argues that the Artificial Anasazi simulation did not give insight into the rules of behavior that governed the Ancient Pueblo people I will argue that the Artificial Anasazi simulation does retain scientific value and it is in the concept of exploratory value.

Grüne-Yanoff and I disagree about the contributions that agent-based simulations can have when developing a 'potential functional' candidate or 'exploratory value' candidate. Grüne-Yanoff suggests that if a simulation is going to make claims of contribution, then it will be important to clarify what sort of contributions the agent-based simulations made and what kind of candidate of explanation it offers (546). According to Grüne-Yanoff, such articulations are going to have shortcomings. Grüne-Yanoff believes that one of the major problems with these kinds of articulations is that when one adjusts parameters within the simulation, whether the adjustment are in landscapes or adjustments to the agent's rules of behavior, those adjustments create major differences to the causal histories that are represented in the simulation.

The problem with this interpretation is that Grüne-Yanoff made the mistake of assuming that the only goal for Epstein and his colleagues was to give a complete account of the causal history of the Ancient Pueblo people, when in fact there were numerous goals for the simulations. One of the goals was for Epstein and his colleagues to work with anthropologists in order to find ethnographically plausible rules of agent behavior" (Epstein, 13). Epstein with the help of anthropologists and the use of ethnographies, they were able to formulate a condensed set of plausible rules of behavior. From that condensed set of rules, Epstein began to select and apply different sets of rules that could be tested in the simulation to see if the agents began to diminish in population (Epstein, 9). This general strategy could be applied to other agent-based simulations where empirical evidence can be used to reduce the number of possible contributions that would cause the system of interest to reach equilibrium. The empirical evidence used to condense the possible contributions, will most likely be part of the empirical evidence used in the building of the model. This general strategy may not give us a complete causal explanation but it may lead us to a better understanding of certain social phenomenon or to further the research and investigation into the system of interest.

One of the many intentions behind the Artificial Anasazi simulation by Epstein and his colleagues was to explore some of the possible factors that may have contributed to the Ancient Pueblo people's disappearing from the Long House Valley. This last point becomes clearer when we take into consideration that Epstein and his colleagues believed that part of the explanation might have been caused by environmental factors. Therefore they knew that through adjustment of rules of behavior, they would only be testing and asserting partial explanation; a clear indication that they were searching for insights rather than causal explanations. Grüne-Yanoff's was mistaking when he assumed that Epstein and his colleagues assumes that researchers in agent-based simulations are looking for causal histories of phenomena. While it could be argued that this may be the final goal of many agent-based simulations, this was not the final goal of the Artificial Anasazi simulation.

The purpose here is not to defend Epstein's and Axtell's claims, I find and argue that they often exaggerate their results. I argue that the purpose of the Artificial Anasazi simulation was not to give us an account of the Ancient Pueblo people's causal histories but instead to give us insight into the different kinds of mechanisms that may have been at play when the Ancient Pueblo people vanished from the Long House Valley. Although I understand that Grüne-Yanoff provides quotes that may imply that Epstein and his collogues where attempting to provide causal histories, there are countless of quotes made throughout numerous publications where Epstein and his colleagues express the idea that, while the simulation was not completely explanatory it was insightful. Epstein undoubtedly supports this last claim when he asserts that the failure of the Artificial Anasazi simulation to "quantitatively replicate the case study results provides valuable insight into what humans might have done in the real Long Valley but did not" (107). This last quote demonstrated the importance of exploratory value, by explaining that the results of the simulation may have failed to give a complete causal explanation, but it was successful in giving insight into what could have been done to avoid such a sudden disappearance. This kind of insight could lead researchers to question why the Ancient Pueblo people did not take steps to avoid complete exodus of the valley.

Epstein understood that generative sufficiency is necessary, but not sufficient condition for explanation." Generative sufficiency, are the micro-specifications that produce the macro-phenomena of interest (Epstein, 53). Epstein knew that even if he was successful at simulating the disappearance, it did not necessarily imply a causal connection.   The point is that agent-based modelers may ultimately strive for complete causal explanations, but they are aware that their simulations may only give insights to the social phenomena of interest.

When a modeler, whether a climate modeler, agent-based modeler or otherwise, are attempting to replicate and simulate a real-world system, they are attempting to attribute some function to their simulations. Again, agent-based modelers are more commonly using their simulations to get insights into further possible lines of inquiry into the target system of interest, illustrating the exploratory value of these simulations.

"From that vantage point, agent-based simulations are but sophisticated ways of formulating hypotheses, and are not in the business of explanation or potential explanation" (Epstein 164). This last quote is just one amongst many, that shows that Epstein and his colleague's main goal was to attempt and identify rules of behavior for the purpose of exploring and theorizing about the possible reasons for the disappearance of Ancient Anasazi, rather than giving a complete causal history. This suggests that the role of the simulation can have a partial contribution to our knowledge to the system of interest.

I argue that the empirical research used in the case of the Artificial Anasazi simulation was a way to pre-select or to 'filter' the potential exploratory possibilities, which in turn negates Grüne-Yanoff idea that the Artificial Anasazi simulation is not supported by direct observation. Grüne-Yanoff argues that, the Artificial Anasazi simulation cannot be supported by direct observation. He believes that the ethnographies used by Epstein and his colleagues were not detailed enough. Grüne-Yanoff cites 'recent research' to illustrate that behavioral rules vary widely among small-scale agricultural societies and because of this, he believes that the differences in behavioral rules can lead to significant changes (544). Those significant changes in behavioral rules can in turn lead to significant changes in the variables that are included in the simulation (ibid). Epstein and his colleagues were able to narrow the infinite number of possible rules of behavior, to rules of behavior that were likely. In other words, there are countless amounts of rules of behavior. For example, different rules for different cultures, different rules for different regions in the world, different rules for large metropolis cities, different rules for small rural communities and the rules of behavior for different communities and societies are extremely varying. Epstein and his colleagues narrowed all of those possible rules, by eliminating those rules that were not characteristic of small-agricultural communities or of other Pueblo tribes.

Epstein, with the help of the ethnographies that he employed, was able to narrow all the possible rules of behavior, to rules of behavior that were likely to be relative, first, to the Ancient Pueblo tribes and second, to people who live in small-scale agricultural societies. So while the rules of behavior in small-scale agricultural societies may vary, more than likely they do not vary in extremes and while rules of behavior may vary from Pueblo tribe to Pueblo tribe, again they do not vary to extremes. What results, are small sets of rules of behavior. Then from those sets of rules of behavior, different rules could be tested in the simulation.

Through induction we can safely assume that the Ancient Pueblo people shared some characteristics with other small-scale agricultural societies, as well as sharing characteristics and traits with other Pueblo tribes. Both, small-scale agricultural societies and Pueblo communities have been directly observed which indicates that we have a categorization of characteristics and traits for each group of people. Then, cross-referenced research of both groups could be used in identifying general characteristics shared by the two groups, which would include characteristics like very simple general rules of behavior. It would be justified to believe that the rules of behavior that were identified, could then be induced as possible rules of behavior for the Ancient Pueblo people.

One of the counter arguments that could be made and supported by Winsberg and Humphreys, is that due to the scarcity of data for the system of interest, one could not confirm or deny any of the results of the simulations. This in turn would suggest that Grüne-Yanoff is correct when arguing that direct observations cannot be a form of support for agent-based simulations. I think that this argument is justified, but only when one is trying to make causal claims about the systems of interest and those claims are supported though the observations and results of a simulation. The use of a simulation is in the strictest sense, not direct observation of the real-world system therefore it would be simply incorrect to claim that one could confirm or deny, undoubtedly, through observations, the causal explanation of a system of interest. That is one of the problems with asserting causal explanations of the real-world system through the support of a simulation. In anticipation of this kind of problem, the exploratory value can be used by researchers, to assign probabilities to the likelihood of *which* micro-specifications are likely and which ones are not. The point is that through the use of exploratory value, we are not attempting to confirm or deny causal explanations. Instead I am suggesting that probabilities of likelihood can be assigned to a real-world system, by direct observations of simulations that incorporate real-world empirical.

Returning to the original argument, the ethnographies used by Epstein may have been vague but they were useful in narrowing down some of the possible rules of behavior of the Ancient Pueblo people. Epstein and his colleagues used the simulation to test their sets of possible rules, indicating that the Artificial Anasazi simulation does have exploratory value because it was used in narrowing the possibilities and testing out those remaining possibilities.

As we will recall, part of a simulation's potential functional analysis or exploratory value, involves pinpointing "functional component(s)" missing in an existing model. As we have witnessed in section 2.3.1, Epstein and his colleagues argue that the Artificial Anasazi simulation is missing a socio-cultural 'pull' component. Epstein contends that the socio-cultural 'pull' component was one of the fundamental reasons for the Ancient Pueblo people abandoning the Long House Valley. Without this functional component, the agents in the simulation never abandon the landscape. In other words, the simulation never reaches the desired equilibrium, thus the simulation was successful in pinpointing the functional component that was missing, namely some socio-cultural factor. On the other hand the simulation was not able give the exact socio-cultural factor that was missing in order for the simulation to reach the phenomenon of interest, which implies that there is more work and research to be done refecting the exploratory value of the simulation.

Grüne-Yanoff believes that agent-based simulations cannot help in the formulation of hypotheses or in giving scientific contributions. In order to formulate possible hypotheses using simulations, according to Grüne-Yanoff, one would be required to identify all the models that simulate the target data and seek to replicate one model's simulation results with another model's simulation results. This is not common practice with researchers in the field, so then according to Grüne-Yanoff, it remains a problem (549). Here Grüne-Yanoff is arguing for the kind of maturing and self-vindicating process advocated by Winsberg and Hacking. If we recall, there is a maturing and self-vindicating process that a simulation must go through in order for the simulation's data to be considered reliable. Each time that the results from a simulation are successful, their credibility as a reliable source of data, increases as well. It should be assumed that this kind of maturing process where "researchers seek to replicate one model's simulation results with another model," is going to be common practice in this field very soon[3]. As more researchers understand the value of agent-based simulations as a scientific tool, there will be a desire and ability to replicate one model's simulation results, with another model, which in turn will help give us possible hypotheses and possible contributing components of macro-phenomena. So even if it is currently not common practice to replicate one model's simulation results, it is not hard to imagine that in the near future this kind of mature, self-vindicating

[3] This would mean publishing code and data sets. This is beginning to happen.

process where replicating a models result with another model, will soon become common practice.

Although currently it may be difficult to get any kind of reliable hypotheses or contributions in social science, from agent-based simulations, it is not impossible I agree that the Artificial Anasazi simulation may not have contributed as a reliable source of data. On the other hand, I do believe that the simulation did contribute to some line of investigation, namely some of the possible rules of behavior of the Ancient Pueblo people.

Perhaps this would be a good opportunity to pause and examine how exploratory value has changed or been supported after the discussion on Grüne-Yanoff's ideas of potential functional analysis. We argued that the functional component missing in a simulation is the kind of insight that demonstrates exploratory value. We also argued that agent based modelers most often are not looking for causal explanations but instead are looking for insights into the mechanisms of the systems of interest illustrating the exploratory value. We have argued that the Artificial Anasazi simulation had exploratory value because further research and investigation are necessary to confirm the possible rules of behavior that Epstein and his colleagues developed. We have also argued that successfully simulating the phenomena of interest does not necessarily constitute a causal connection. We showed how Epstein successfully used ethnographies to narrow the possible rules of behavior and tested those possibilities in his simulation, another indication of the Artificial Anasazi simulation's exploratory value. In the following chapter we will continue to look at how the results of a simulation contribute to the exploratory value of a simulation.

Chapter 4: Exploratory Value and Potential Functional Analysis

I will briefly review Grüne-Yanoff ideas on what he believes are sources of evidential support for simulations, then I will continue reviewing some of his idea on how the quality of a simulation's functional analysis can be assessed by the formal properties of the simulation. Then, I will argue that evidential support will include the simulation itself. Toward the end of chapter I will introduce a discussion concerning how computational science and computational simulations will force a reevaluation of epistemology of philosophy science. I will argue that in order for philosophers to engage the philosophy behind computer simulations then philosophers will need to become concerned with problems of application. In turn, these new for philosophers considerations are an indication that the use of agent-based models will require an interdisciplinary efforts.

Grüne-Yanoff argues that, evidential support for a simulation cannot come from the simulation itself and that there are three sources which can potentially be sources of support for simulations, which are: direct observation, well-confirmed theory, or results from externally valid behavior experiments (544). Grüne-Yanoff argues that the Artificial Anasazi simulation does not have any of the three potential sources of support, so then it cannot give any causal explanations

Of the three forms of evidential support, I believe that I have sufficiently argued that direct observations cannot be performed on the Ancient Pueblo people because they have stopped existing. Nonetheless, direct observations from other very similar groups of people are well documented and therefore a cross-reference of these groups of people could successfully be applied to identify traits that are likely to be analogous to the Ancient Pueblo people.

The second source of support that the Artificial Anasazi simulation lacks, according to

Grüne-Yanoff, is behavior experiments. Grüne-Yanoff argues that the Ancient Anasazi went through four distinct periods of distress. He argues that the four distinct periods of distress must have had an effect on the rules of behavior of the Ancient Pueblo people. And because the agents in the Artificial Anasazi simulation were commanded to behave the same throughout the four distinct periods, then the data produced by the simulation is not transferable to the Ancient Pueblo peoples case. I agree with Grüne-Yanoff that the rules of behavior would have most likely changed with the different distinct periods, but this was a problem of application and not of principle. In principle, when each of the distinct periods change, then the rules of behavior could be instructed to change as well, this is one of the distinct features of agent-based models and discussed in more detail in chapter 2. This observation by Grüne-Yanoff, is the kind of observation that needs to be accounted for, implemented and tested in the Artificial Anasazi simulation.

Agent-based simulations are capable simulating dynamic change, which in turns suggest that the four distinct distress periods, that Grüne-Yanoff describes can be applied to the Artificial Anasazi simulation. As more researchers construct more models, more attention will be given to particular models, more recommendations and critiques will in turn lead to adjustments of the simulation that will then produce better more reliable results. It is a process of maturing that I reviewed earlier and will go further into detail that I believe gives agent-based simulations exploratory value

The final source of support for simulations as suggested by Grüne-Yanoff is well-formed theories. The objection here is reasoned much like the objection discussed earlier for direct observation. Grüne-Yanoff argues that there is no well-formed theory of behavior regularities, therefore there is not decisive evidence for behavior regularities required for artificial societies.

It is quite obvious that there is not going to be any grand overarching theory of behavior regularities because behavior regularities are going to be partially influenced by ones time in history, cultural and social background. Cultures vary in behavior regularities and thus, it will be extremely difficult, if not in possible to make any kind of generality about behavior regularities. We are not necessarily trying to identify behavior regularities for all cultures. If one were to attempt to identify rules of behavior, it would most likely be for a particular kind of people or culture. Therefore an overarching theory of behavior that includes all groups of people is not necessary. Through the use of ethnographies, statistics, and other forms of data and methods developed by anthropology and sociology, one could identify and develop theories of behavior regularities of a certain groups of people. This is what Epstein and his colleagues attempted to do. It would seem as if they narrowed the scope enough so that the theoretical rules of behavior, developed for the agents in the simulation, were not overly general. The reason why the rules of behavior were not overly general is because they were developed from cultures that were very similar and whose nutritional requirements and number of members were relatively similar as well. It is understand that there is no general theory of behavior regularities and perhaps there never will be. We can develop general ideas on relatively small communities, which would provide decisive evidence for behavior regularities. Then those behavior regularities could be used in the simulations of artificial societies. The point is that evidential support for simulations must include the simulation. Data obtained from direct observation, well-confirmed theory, or results from externally valid behavior experiments, should be incorporated as much as possible into the simulation.

Although production and analysis of a simulation may not be the evidential support that Grüne-Yanoff had in mind, it is still imperative that the structure and organization of the agent-based simulation be taken into account when assessing a simulation's credibility. If a simulations' structure and other formal properties are of poor quality, then the simulation will have results that cannot be trusted. On the other hand, if we find that a simulations' structure, organization and other formal properties are intact, then the those properties should be considered part of the whole process that leads us to trust the simulation.

Organization, structure and other formal properties may not have been traditionally considered evidential support but in agent-based simulations these kinds of properties can be and often are the empirical evidence used in the simulation. As we recall, from the Winsberg discussion, it is the combination of proper model building techniques, our prior intuitions, empirical data along with other prior successes, matched with the proper background knowledge of the target system and the proper knowledge of the "principles of modeling," that ultimately gives us confidence in the model and its results (Winsberg, 35). It will be crucial that researchers in the field take precaution in constructing each simulation, so that when our simulations fail at giving causal accounts, we may have an idea where the problem lies

Of course it would be a mistake to assume that an agent-based simulation can become self-validating and mature enough to give us reliable data, by simply examining the simulations' formal properties. We should recall that in order for a simulation to give us reliable data, we would have to trust the incorporated empirical evidence. The significant point is that the combination of proper model building techniques and our prior intuitions and data that will eventually lead agent-based simulations to be mature and self-vindicating enough to be reliable sources of data. As Winsberg stated, "it is the simultaneous confluence of these efforts, rather than the establishment of each one separately that ultimately gives us confidence in the result"

(23). So in order for simulations results to be trusted, then evidential support is required from several components of the simulation.

4.1 Conclusion on Exploratory Value

Exploratory value is assigned when an agent-based simulation has failed in giving researchers a complete causal explanation of their phenomena interest but has given insight into the phenomena of interest. The Artificial Anasazi simulation obviously cannot replace archeological evidence for several reasons. The first reason is because the simulation has not had the necessary time to evolve and mature enough. On the other hand, the simulation matured some, in the sense that adjustments were necessary and new lines of research resulted, which is a reason the simulation retained exploratory value. The second reason that results from the Artificial Anasazi simulation cannot replace archaeological results is because of the simulation's inability to be self-vindicating. The simulation does not have the history of prior success that warrants reliability and trustworthiness. Reliability and trustworthiness of casual explanations are also not required in simulation in order for it to retain exploratory value.   The Artificial Anasazi simulation did reproduce some of the aggregated and dispersal patterns and because of this, some of the rules of behavior embedded in the simulations may be correct. In turn the simulation has given us possible insight into what may have been partial rules of behavior that caused the Ancient Pueblo people to abandon the Long House Valley. Giving possible insight into some of the potential rules of behavior is the kind of insights that epitomizes explanatory value.

If we are inclined to believe Humphreys assertion that computer models and other forms of technology are a way of extending ourselves and enhancing our natural epistemic abilities, then it would be safe to say that the Artificial Anasazi simulation does have some scientific value and the value derives from the exploratory nature of the simulation. Insights gained from the simulation, such as the 'functional component' missing, will allow for further investigation into the possible reasons of the abandonment of the Long House Valley. Although we may never have a complete causal account of the social factors that led to the abandonment, we can gain a better understanding by testing social rules in the simulation and then eliminating the rules that do not lead to the target equilibrium

Exploratory value constitutes the type of insights that agent-based modelers are interested in, rather than expecting their simulations to explain the complete causal accounts of the phenomena of interest. Evidential support for a simulation will include outside sources such as direct observation, well-trusted theories as well as an analysis of the internal structure and formal properties of the simulation.

In the following section, with the aid of Humphreys's argument on syntax and semantics, we will understand why context will be important in simulations and will conclude that the context will push philosophers to concern themselves with nontraditional problems of application.

4.2 Introductory Discussion of Epistemic Changes and the Emphasis on Interdisciplinary Efforts.

In this section I give a brief introduction to Roman Frigg's and Julian Reiss's response to philosophers who have claimed that computer science and computer simulations demand a new epistemology of philosophy of science. These critiques are particularly useful because two of the philosophers addressed by Frigg and Reiss, are presented in this thesis, namely Winsberg and Humphreys. I will briefly introduce Frigg and Reiss's responses, which they have divided into four parts: Metaphysical, Epistemic, Semantic and Methodological. I will briefly respond to

Frigg and Reiss in order to facilitate a discussion on the denial that there are new epistemic problems that are particular to computer simulations.

One epistemic problem that is particular to computer science is the need for interdisciplinary work in order for agent-based models to be constructed, realized, adjusted, tested and analyzed. In the Metaphysical section of their response, Frigg and Reiss at times state, and at times imply, that there are two claims being made by those who advocate that parallel worlds are unique to simulations. Frigg and Reiss are arguing that artificially created system are used all over in science. Frigg and Reiss use the discussion on artificially created systems in order to imply that the artificial systems are often deemed favorable over the real-world systems, when it comes to experimentations. I do not believe that artificially created systems are favored over the real-world systems because it would seem that if it was physically possible or ethically valid, most scientists would prefer to experiment on the actual real world systems itself. Scientists are aware of the problems of induction because their work is intimately tied with experiments and applications. They understand the problems of inducing the explanation of one phenomenon as the explanation for another phenomenon. I presume that if scientists could avoid these kinds of problems of induction by experimenting and observing the real-world system rather than a proxy system, they would. If researchers could test and manipulate the real-world system, it would save them from the exhausting efforts of creating a proxy system, testing that proxy system and then interpreting those results. The problem of inferring the results of the proxy system to the real system of interest would cease to exist.

As previously mentioned, researchers for whatever reason do not have access to the real-world systems, making experimentation difficult if not implausible in some cases. So what simulations do, as advocated by Humphreys, is that they extend our abilities to investigate phenomena, which we would otherwise not be able to investigate. That a simulation creates a parallel world that is favorable to the real world is not a correct assumption by Frigg and Reiss. The real-world system in most cases, if not in all, would be favored since we do not have physical, cognitive and ethical access to certain systems. We do not have physical access to systems such as astronomical bodies. We also do not have complete cognitive access in cases where the amount and extent of calculations being done by a computer go beyond our abilities to track. Finally, we often do not have ethical access to systems, like in the cases of medical experimentation on humans. The only possible choice, when attempting to investigate systems such as the ones just mentioned, other than discontinuing investigations all together, is to create parallel worlds. It is quite clear that the results of experimenting on real-world systems would be more reliable and favored over the results produced by proxy systems. Therefore I do not believe that most researchers would favor proxy systems over real systems. The problem lies in our access and ability to manipulate those systems, if it is not possible, then the creation of parallel worlds will have to do.

The second claim made by Frigg and Reiss, is that the creation of parallel worlds does not require a new philosophy of science (598). I would argue that Frigg and Reiss are correct when asserting that creating parallel worlds are not distinct to simulations. On the other hand, I would add that the problem of the amount of detail of approximation, between the proximate system and the system of interest is unique to simulations. To illustrate this point we can use the same examples Frigg and Reiss's used. They argue that we create parallel worlds when:

> "We infer from a small group of patients in a clinical trial to the general population; we experiment with mice to find out whether smoking causes cancer in humans; we examine the properties of a scale model of an aeroplane wing in a wind tunnel to learn about the properties of the aeroplane wing during a flight and so forth."

The parallel worlds described above, do not have many details in common with the real-world systems. At times, this lack of details may not hinder experimenting but at other times the lack of details may have disastrous consequences such as, in cases where results in medical experimentations on animals, are applied to human health problems. Experimental in data that is produced by performing experiments on animals is hardly ever transferable to humans; their biological systems are too fundamentally different. If it was ethically permitted researchers would get a lot more reliable results if they experimented on humans but because they cannot they must use proxy systems i.e. other animals, those proxy systems are often much too different in details to assume results from one system to another. The point is that researchers would benefit from using the actual systems of interest but because they cannot, researchers must find proxy systems whose results in may not be transferable.

When we create parallel words, we are attempting to recreate the system of interest with as much approximation to details as possible. After all, the goal is usually to get a better understanding of the phenomena created by the real world. One problem with researching large complex dynamical systems is that those systems create an overwhelming amount of transactions and relationships that become nearly impossible to track without the aid of instruments. Simulations can help by allowing us to input large amounts of data into the simulation and then keeping track of the relationships and interactions displayed by the simulation. This is what Humphreys refers to as “detail of application” and it is a specific problem in computer science, according to Humphreys. As I argued before, if scientists had the necessary access and abilities to experiment on the real-world system, most likely, they would. However, since scientists often do not have access to real-world a system, the next best option is a proxy system that replicates as much relative details as possible. As our technological achievements progress, I suspect that the amount of detailed data that can be replicated, tested and analyzed, will increase. In turn, one problem that agent-based modelers will have will be to distinguish how much or what kind of detail they will have to incorporate into their simulations in order to be able to accurately study the system of interest. Frigg and Reiss were wrong to assume that there are not any new epistemic problems that are unique to computer simulations.

While I would not defend a statement such as, “the creation of a parallel world is reason to believe that simulations require a new philosophy,” I would on the other hand, defend a statement that suggests that simulations create new philosophical problems, that are particular to computer science. New philosophical questions that are particular to computer science would include questions such as, what kinds of details are sufficient and how much details are sufficient to create a simulation that replicates the system of interest. As Humphreys suggests “there is now a growing sense that a different problem has arisen: that new techniques need to be developed to effectively exploit the massive computational power that is now available in many area” (2009, 18).

Frigg and Reiss deny Winsberg’s assertion that there are three features that make the epistemology of simulations distinct. The three features are downward, autonomous and motley and thus, Frigg and Reiss argue that none of these features are distinct to simulations and can all be found in other sciences.

Frigg and Reiss then claim that a different reading of Winsberg suggests that the results from simulations and computational models are justified, at least in part, on the basis of the principles and techniques used in the construction of the computational model. Winsberg and Humphreys argue that the result of the simulation cannot be dismissed as unjustified because they cannot be compared directly with observational data and the target system. The results of a simulation are justified, in part, on the basis of approximation, simplification, idealization and isolation. Frigg and Reiss argue that none of these features, described by Winsberg and Humphreys, are unique to simulations. They believe that the same conclusion could have been reached by studying the practice of other sciences. Frigg and Reiss claim that the need for a new epistemology is due "to the fact that models are more complex than traditional philosophy of science allows and that we still do not have a worked out epistemology that accounts for this". They argue that although novel and specific questions come out of the use of simulations, those questions belong to the class of problem that arise in connection with complex models and in general. Those problems that arise in connection with complex models according to Frigg and Reiss, "do not represent a revolutionary departure from everything that philosophers were worried about."

I do not assume that the epistemology that arises from computational simulations is "a revolutionary departure from everything that philosophers are worried about" (601) as Frigg and Reiss suggest that some philosophers have claimed. On the other hand, I would agree with Frigg and Reiss' claim that a need for a new epistemology is due "to the fact that models are more complex than the traditional philosophy of science allows and that we still do not have a worked out epistemology that accounts for this…specific issues arise when we look into the details of computational models and agree that part of the justification of the models comes from the fact that they are based on theory that is well understood" (601). It has become evident through the research of computational models that science is slowly reaching the limits of what can be known through our unaided, cognitive abilities. This is clearly the case that Paul Humphreys is making in his book, Extending Ourselves. As discussed before, Humphreys argues that the nature of epistemology is changing and has moved from what Humphreys refers to as human based epistemology to a technology based epistemology. He suggests that one of the principle achievements of science has been to transcend the limitations of the humans' natural epistemic cognitive abilities in order to investigate complex phenomena. Different fields in science are increasingly recognizing that an epistemology that is based on human cognitive capabilities is no longer appropriate because there are non-human instruments and other forms of technology that are now necessary in forming the base knowledge of our sciences. In turn a new problem must be dealt with, one that is distinct and specific to computer science. The problem is referred to as, *anthropocentric predicament*, by Humphreys. He elegantly summarizes the problem of *anthropocentric predicament* when he states:

> "This predicament is different from the traditional philosophical problem of understanding the world from a human perspective because the traditional problem involves representational intermediaries that are tailored to human cognitive capacities"(2009, 3).

Part of the answer to the problem of anthropocentric predicament is found when we combine the need for human use, with the need for computational tools. As Humphreys observes, Frigg and Reiss never deal with important aspects of computational science even though this aspect reflects a key change in epistemology. Aspects in computational science have consequences that are philosophical, i.e. how we think and view models, theories and other representational tools (2009,4).

In the section about Syntax and Semantics, Frigg and Reiss contend that:

> "In the broadest sense of application-meaning simply the entire process of using the model, we use computational methods rather than paper and pencil to get the solutions of the equations that form part of the model…then this is just a restatement in "application jargon" that there are equations which defy analytical methods and have to be solved numerically" (604).

The problem here is that it is not always very clear or easy to apply computational methods in order to solve equations, which have to be solved numerically (2009,9). The difficulty of application tasks in computer science are much more complex than Frigg and Reiss imply, and according to Humphreys, issues concerning, "opacity, dynamics and possibility in practice" are "relevant to the process of computationally applying a scientific representation to a real-world system"(ibid).

Frigg and Reiss find it puzzling that Humphreys states that, "neither syntactic or semantic theories, models or research programs and paradigms, capture what simulations can do." Frigg and Reiss contend that, simulations by themselves do not clash with the semantic or the syntactic view (606). Humphreys argues that neither syntax nor semantics are the proper vehicles to represent computer science (2009,9). Syntax, according to Humphreys, cannot be separated from the actual use of simulations because simulations are implemented in actual time and it involves processing on machines (2009,10). Humphreys argues that, the only way to realize the simulation is to run the code. Humphreys maintains that the semantic view of theories abstracts from language used in the theory of interest and identifies theory with its class of …set theoretical structures that make the various linguistic formulations true.

The problem with the syntactic and semantic approach arises when questions like, how *d*o we apply theories are brought up. The act of abstraction that is necessary in the semantic approach, limits the differences that are crucial in applying theory. It is clear that Humphreys is essentially arguing that the context construction and data implemented in simulations will need to be specific to the system of interest. If this claim were true, then it would not be difficult to understand the essential move that will be required of philosophers in order to continue commenting in the philosophy of science. Since researchers are often looking for a connection between theory and application in computer science, philosophers will need to concern themselves with problems in practice rather than just problems in principle. This is clearly what Humphreys means when he states that computational science suggests that a less abstract approach to scientific method should be used by philosophers (2009, 12). If philosophers fail to concern themselves with the practical application of theory, then their comprehension and commentary on the subject of computational modeling and computer simulations will be lacking as well. Philosophers are not required to concern themselves with the practical application of simulations, but their narration on how computer science works and progresses will have shortcomings. This last idea is nicely summarized when Humphreys states that "ignoring implementation constraints can lead to inadvisable remarks" (2009-15).

In the section titled: Methodology: does simulating constitute a Sui Generis activity, in between theorizing and experimenting? Frigg and Reiss argue, that the phrase "falls between" or "lies between" can be taken in a literal sense or metaphorical sense. I assume that Winsberg meant to be interpreted in a metaphorical sense. Therefore I will limit my response to Frigg and Reiss's objections to the metaphorical interpretation. There are, according to Frigg and Reiss, important classes of methods in science that are "in between" theorizing and experiments in the same way but that have nothing to do with computers (ibid). They conclude that philosophical problems that are raised by simulations have analogies in modeling, experiments and thought experiments, which in turn implies that those problem are not new and do not demand a novel philosophy of science (611).

The best way of interpreting the phrase "lies in between" is metaphorically, this is clearly how Winsberg's intended to uses the phrase. He suggests that the reason why simulations are between theory and experiments is because simulations, much like experiments are credited and become reliable through the process of maturing. Winsberg offers a helpful illustration of this idea when he quotes Deb Dowling's observations, that simulations are like theory in that it involves "manipulating equations", "developing ideas", and simulations are like experiments because both involve "fiddling with machines" and "watching what happens" (90). Simulations, like experiments depend on the quality and background knowledge, in which the simulation is implemented, realized, tested, adjusted, and analyzed. According to Winsberg, with both simulations and experiments, you need to know something in order to learn something (71). The major difference between simulations and experiments is the background knowledge that is required in simulations, according to Winsberg, "is always quite abstract and sophisticated and usually depends on what you have learned in a long history of experiments and observations" (ibid).

An example of the abstract and sophisticated knowledge needed in creating an agent-based model is the Artificial Anasazi simulation. In this thesis the creators of the Artificial Anasazi simulation are consistently referred to as "Epstein and his colleagues ", this is bit unjust because there were extensive amounts of work, by numerous researchers and scientists, throughout numerous fields that was implemented and the in the development of the simulation. Epstein in his book makes great efforts to include and reference all of the research and individuals that were involved with the simulation. The steps required to research, theorize and apply the simulations, is an impressive demonstration of the interdisciplinary work that was necessary for the simulation. In order for Epstein to make the adjustments, that are relevant in making his simulation reach the target phenomena, more research, and experts with different specialized knowledge, would be needed. As the system under investigation increases in complexity, it will be no surprise to expect, that the amounts of data and experts, required in the simulation of those complex systems, will increase as well

Humphreys has this last idea in mind when he suggests that computational science made the field of Complexity possible. Complexity, being a new science has it own methods, powerful interdisciplinary capabilities. The methods involved in complexity are different from those taught in theory classes and in laboratory sessions, illustrating that the use of computer simulations are in fact, between theory and experiments (2009,19).

Frigg and Reiss conclude by restating their argument, which is that features such as idealization, approximation simplification and isolation, are not specific to computer science. While this may be true, Frigg and Reiss forget to address problems that are specific to computational models. Problems that are specific to computational science require philosophers to reevaluate the epistemology of science when they are faced with those new and specific problems. Humphreys articulates some of the problems that are specific to computer science when he states that the problems include

> "Constraints put on models by computational load issues, the problems of extending models when substantial chunks of existing code are written in legacy software, the choice of finite element decomposition, and the need for research teams to delegate substantial amounts of authority to programmers"(2009,17).

There are new philosophical concerns that emerge from computer science and these new concerns should persuade philosophers to do work that is fundamentally different and regards the problem of practice and application of the epistemic approach in the philosophy of science.

Another indication that a reevaluation of the epistemology of philosophy is imminent is one that is hinted at throughout this thesis, namely, the interdisciplinary work that will be vital in agent-based simulations. If we notice, simulations require abstract and sophisticated knowledge, as suggested by Winsberg. In order to interpret, analyze and implement such abstract and sophisticated knowledge, efforts will need to be made across disciplines. Philosophers will no longer be limited to problems in principle, but will need to confront problems that come with the practice and use of simulations. If philosophers plan for their work to be significant, then they will be persuaded to concern themselves with problems of applications in computational science. Winsberg supports this idea when he claims "philosophers have missed the opportunity to contribute to science because they have prejudice for asking and concerning themselves with problems that are possible or not in principle rather than concerning themselves with what could be achieved in practice" (7). The epistemology of simulations is intimately tied to the application of scientific theory, which in turn suggests that philosophers are relatively unfamiliar to this new approach in epistemology. Accordingly the relationship between philosophers and science will also need to be reexamined. A scientific work that can be done within one field or discipline has changed from individuals to group projects; it would seem that another change is in order to continue scientific investigations

It is evident that due to the fact that epistemology is changing from a human based epistemology to a technology based epistemology, philosophers in order to continue commenting on the philosophy of science, will have to understand the use and application of the technological instruments that are becoming fundamental to science. One implication of this move from theory to application, suggests that the efforts that will be necessary to construct, test, run, adjust, and analyze will include information from various and distinct academic fields. The complex systems which are most often the subjects of simulations require extensive amounts of data in order for that data to be properly chosen, analyzed and implemented into a computer simulation, efforts from different occupations will be necessary. This is supported when Mitchell suggests that the topics concerning complex systems require an interdisciplinary field of research. This is also evident when Humphreys suggests that there are parallels with the switch from individualistic epistemology to a social epistemology because no single scientist or mathematician can properly verify the procedures and proofs needed in computer simulations (2009,4).

We are reaching the limits of what our specialized sciences can tell us about the phenomena in the world and are slowly being forced to look at how relatively simple agents function and relate to one another in order to create complex adaptive behavior(s). Mitchell suggests that complex, dynamic, social phenomena cannot be pigeonholed into any single discipline but require an interdisciplinary understanding and can these kinds of social phenomena can only be studied properly in an interdisciplinary environment (x).

Chapter 5: Conclusion

I have introduced the distinctive value of agent-based models, which is called the exploratory value and attempted to illustrate how the exploratory value may benefit the social scientists. The clear benefit of agent-based models for social scientists may have been muddled by the responses to the philosophers. It would seem appropriate to give a clear and efficient description of the value of agent-based models to social scientists. Exploratory work can (1) from a partially failed model to raise questions about what functional components might need to be added; (2) from a relatively successful model, raise questions and continue investigation into what functional components might be at work; (3) continued inquiry into complex dynamic system that would otherwise be unfeasible to inquire into; (4) continued experimentations by allowing other researchers to replicate and adjust models; (5) confirm a model on a robust data set, and then extend it to a place where the confirmatory evidence is more incomplete;(6) the possibility of confirming the results of models by other researchers; (7) simulate the dynamic change, that a society and culture may experience (8) experiments with counter-intuitive possibilities (9). These are a few of the ways that I have argued for the exploratory value of agent-based models for social scientists.

Social science often deals with some of the more pressing issues faced by humans, such as the spread of diseases, the unequal distribution of resources and the affects of our societies and cultures on the environment. Agent-based models offer another method of inquiry and experimentation into the complex systems that are involved in these kinds of pressing issues. Issues like the ones given above are not issues or concerns of any single discipline and in fact different aspects of these issues are already investigated by different disciplines. It would appropriate that combined interdisciplinary efforts are made to further investigate the systems that are involved in these issues.

Agent-based models are a new approach where combined interdisciplinary efforts and current computational power, gives us insights into systems that have been to complex to properly investigate. Because of the computational science involved in the new approach, the agent-based models will have new issues that are specific to computational science. These new issues will force, not only social scientists and other researchers interested in complex system to reconsider methods in their discipline, it will also force philosophers to revaluate the epistemology of science. I have argued that reevaluation will imply that philosophers will need to concern themselves issues of applications.

Agent-based models, will continue the scientific exploration into complex, dynamic systems that were often too complex to be investigated by any single discipline and because of this, philosophers will need to be on the foundational ground in order to provide analysis and clarity into the new ideas and terms that will be needed in the continued research. The hope is that enough arguments have been provided to support the claim that agent-based models will be an invaluable tool to science in general and social science in particular and.

List of References

Epstein, Joshua. M. *Generative Social Science: Studies in Agent-Based Computational Modeling.* Princeton, NJ: Princeton University Press, 2006. Print

Frigg, R. & Reiss, J. "The Philosophy of Simulation: Hot New Issue or Same Old Stew?" *Synthese* 169. No 3. (2009): 593-613. Print.

Grüne-Yanoff, Till. "The Explanatory Potential of Artificial Societies." *Synthese* 169. No 3. (2009): 539-555. Print

Hacking, Ian. "Do Thought Experiments Have a Life of Their Own?" *Proceedings of the 1992 Biennial Meetings of the Philosophy of Science Association East Lansing, MI*. Ed. A. Fine, M. Forbes, and K. Okruhlik. East Lansing, MI: The Philosophy of Science Association, 1992. Print.

Humphreys, Paul. *Extending Ourselves: Computational Science, Empiricism and Scientific Method*. New York, NY: Oxford University Press, 2004. Print

______. "The Philosophical Novelty of Computer Simulation Methods." *Synthese,* 169. No. 3. (2009): 615-626. Print

Mitchell, Melanie. *Complexity: A GuidedTtour*. New York, NY: Oxford University Press, 2009. Print

Voss, Jesse. L. "Scientific Grounding for Complex Agent Based Model Development, Reverse Engineering, and Redesign: Use of NVivo + UML on the Long House Valley Artificial Anasazi Model as a Round Trip Data Tracking Protocol." *CSAAW 2009 Spring Conference Ann Arbor, MI*. Ed. Center for the Study of Complex Systems. Ann Arbor, MI: University of Michigan, 2008. Print

Winsberg, Eric. *Science in the Age of Computer Simulation*. Chicago, IL: The University of Chicago Press, 2010. Print

Curriculum Vita

Ricardo Rivera was born and raised in El Paso, Texas. He is a graduate of Del Valle High School, Class of 2000. Ricardo began his Undergraduate degree at The University of Texas at El Paso in the Fall of 2004. Prior to entering UTEP, Ricardo took a philosophy class at a community college and that class peaked his interest in philosophy. Ricardo has received a B.A. in Anthropology and Philosophy. He hopes to continue to study philosophy of science.

Permanent Address: 8944 Sandhill Ct.
El Paso TX.
79907