McLaughlin, Brian P. (1992). "The Rise and Fall of British Emergentism." In *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism,* eds. Ansgar Beckerman, Hans Flohr, and Jaegwon Kim, 49–93. Berlin: Walter de Gruyter.
Mitchell, Melanie (2009). *Complexity: A Guided Tour.* New York: Oxford University Press.
Morowitz, Harold J. (2002). *The Emergence of Everything: How the World Became Complex.* New York: Oxford University Press.
Mouritsen, Ole G., and Ask F. Jakobsen (2008). "Forming the Essential Template for Life: The Physics of Lipid Self-Assembly. In *Protocells: Bridging Nonliving and Living Matter,* eds. Steen Rasmussen, Mark A. Bedau, Liaohai Chen, David Deamer, David C. Krakauer, Norman H. Packard, and Peter F. Stadler, 385–406. Cambridge, MA: MIT Press.
Nagel, Kai, and Steen Rasmussen (1996). "Particle Hopping Models and Traffic Flow Theory." *Physical Review E* 53: 4655–4672.
Packard, Norman (1989). "Intrinsic Adaptation in a Simple Model for Evolution." In *Artificial Life,* ed. Christopher G. Langton, 141–155. Redwood City, CA: Addison-Wesley.
Rueger, Alexander (2000a). "Robust Supervenience and Emergence." *Philosophy of Science* 67: 466–489.
Rueger, Alexander (2000b). "Physical Emergence, Diachronic and Synchronic." *Synthese* 124: 297–322.
Sawyer, R. Keith (2005). *Social Emergence: Societies as Complex Systems.* New York: Cambridge University Press.
Silberstein, Michael, and John McGeever (1999). "The Search for Ontological Emergence." *Philosophical Quarterly* 49:182–200.
Stephan, Achim (2006). "The Dual Role of 'Emergence' in the Philosophy of Mind and in Cognitive Science." *Synthese* 151: 485–498.
Stano, Pasquale, Giovanni Murtas, and Pier Luigi Luisi, (2008). "Semisynthetic Minimal Cells: New Advancements and Perspectives." In *Protocells: Bridging Nonliving and Living Matter,* eds. Steen Rasmussen, Mark A. Bedau, Liaohai Chen, David Deamer, David C. Krakauer, Norman. H. Packard, and Peter F. Stadler, 71–100. Cambridge, MA: MIT Press.
Sugiyama, Yuki, Minoru Fukui, Macoto Kikuchi, Katsuya Hasebe, Akihiro Nakayama, Katsuhiro Nishinari, Shin-ichi Tadaki, and Satoshi Yukawa (2008). Traffic Jams without Bottlenecks: Experimental Evidence for the Physical Mechanism of the Formation of a Jam. *New Journal of Physics* 10. doi:10.1088/1367–2630/10/3/033001.
Weinberg, Steven (1987). "Newtonianism, Reductionism, and the Art of Congressional Testimony." *Science* 330: 433–437. (Reprinted in eds. Mark A. Bedau and Paul Humphreys, 345–357, 2008).
Wimsatt, William C. (1986). "Forms of Aggregativity." In *Human Nature and Natural Knowledge,* eds. Alan Donagan, Anthony N. Perovich, Jr., and Michael V. Wedin, 259–291. Dordrecht: Reidel.
Wimsatt, William C. (1997). "Aggregativity: Reductive Heuristics for Finding Emergence." *Philosophy of Science* 64(4): 372–384.
Wolfram, Stephen (1985). "Undecidability and Intractability in Theoretical Physics." *Physical Review Letters* 54: 735–738.
Wolfram, Stephen (2002). *A New Kind of Science.* Champaign, IL: Wolfram Media.

6 Agent-Based Modeling and the Fallacies of Individualism

Brian Epstein

Agent-based modeling is starting to crack problems that have resisted treatment by analytical methods. Many of these are in the physical and biological sciences, such as the growth of viruses in organisms, flocking and migration patterns, and models of neural interaction. In the social sciences, agent-based models have had success in such areas as modeling epidemics, traffic patterns, and the dynamics of battlefields. And in recent years, the methodology has begun to be applied to economics, simulating such phenomena as energy markets and the design of auctions.[1]

In this paper, I aim to bring out some fundamental limitations and tradeoffs to agent-based modeling in the social sciences in particular. Two misconceptions about social ontology, pertaining to the relation between social macro-properties and individualistic properties, are widespread in social theory and modeling. These issues lead current models to systematically ignore factors that may be significant for modeling macro-properties. To treat the problem, I suggest that we give up on two deeply held assumptions: first, that agent-based (and other) models can provide the microfoundations for macro-properties and second, that models have to avoid ontological redundancy. Abandoning each of these is painful but may be less costly than the alternative.

AGENT-BASED MODELING IN THE SOCIAL SCIENCES

In the social sciences, as in the natural sciences, most mathematical models are "analytical" models rather than computational ones. Analytical models typically consist of systems of differential equations, giving structural relationships between variables of interest, such as the Lotka-Volterra equations, describing predator-prey dynamics, and the Susceptible Exposed Infectious Recovered (SEIR) model of epidemic propagation.[2] The most familiar agent-based models are cellular automata, with agents represented as states on a fixed geographical grid. The "Sugarscape" model, introduced by Joshua Epstein and Robert Axtell, is a well-known example.[3] With agent-based models, it is easy to construct a large population of heterogeneous agents. Sugarscape, for instance, represents a population of individuals who may be young, middle-aged, or elderly, who can "see" only nearby cells and who can see distant cells, and who have slow or fast metabolisms. Interestingly, it also includes environmental resources in the model: occupying the cells in addition to people are quantities of "sugar" and "spice," with which the agents interact.

Agent-based modeling is typically understood to be a way to provide micro-foundations for changes in macroscopic properties. Consider some macroentity, say France, represented in Figure 6.1 as a hexagon.

France has a variety of macro-properties, such as its inflation rate, unemployment, and monetary financial institution interest rates. One way to model the interrelations among these variables is to come up with macrolaws (L_M) that govern the relationships. In macroeconomics, the values of parameters might be determined by statistical estimation. A micro-foundational model attempts to decompose and eliminate the macro-properties. Instead of modeling the macro-properties themselves, it models interactions among individual agents, each of whom starts out in a particular state, with the system updating over time, as shown in Figure 6.2. At any time-step, the macro-properties can be "read off" of the micro-states.

Many analytic micro-foundational models have been developed in economics, but they tend to involve a "representative agent" framework, in which individuals are represented as if the aggregate of their choices is equivalent to the decision of an aggregate of identical individuals.[4] Agent-based modeling allows more complex initial states and transition rules to be incorporated.

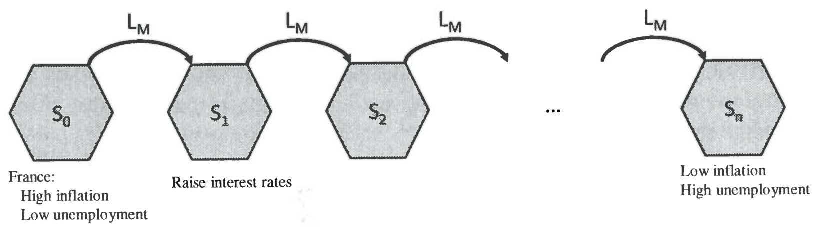


Figure 6.1 Macroentity represented as a hexagon.

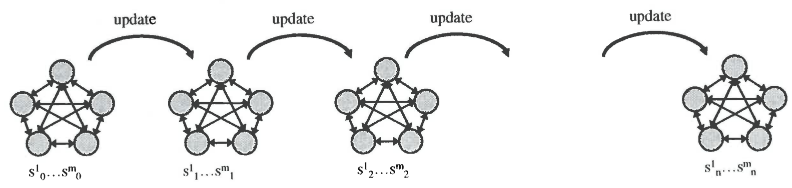


Figure 6.2 Interaction among individual agents.

Although the best-known agent-based models are cellular automata, there are many different types, with more general mathematical characteristics. Cellular automata involve changing states of a fixed homogeneous set of cells that neighbor one another. They also typically involve a synchronous updating schedule, all the cells being updated at the same time.[5] In a more general agent-based model, one can choose to model very different sets of objects, apart from locations or people, and agents do not need to be modeled on a grid. An agent-based model, for instance, might involve a network of power plants interconnected by power lines. The network configuration itself can also be dynamic. In a model of the evolution of an epidemic, people move over time, so the network of interactions changes from one moment to the next. And different agent-based models involve different updating schedules, so that agents can update and interact asynchronously in various ways. Finally, processes in the system can be stochastic as well as deterministic.

Generically, these can be described by what might be called "graph dynamical systems," mathematical structures that include cellular automata and other structures as subcases.[6] One simple and very general class of structures, widely discussed in connection with agent-based modeling, is the sequential dynamical systems.[7] A typical agent-based model starts with a set of objects or agents, each in a certain state and with a network of connections among them. In a sequential dynamical system, this is represented as

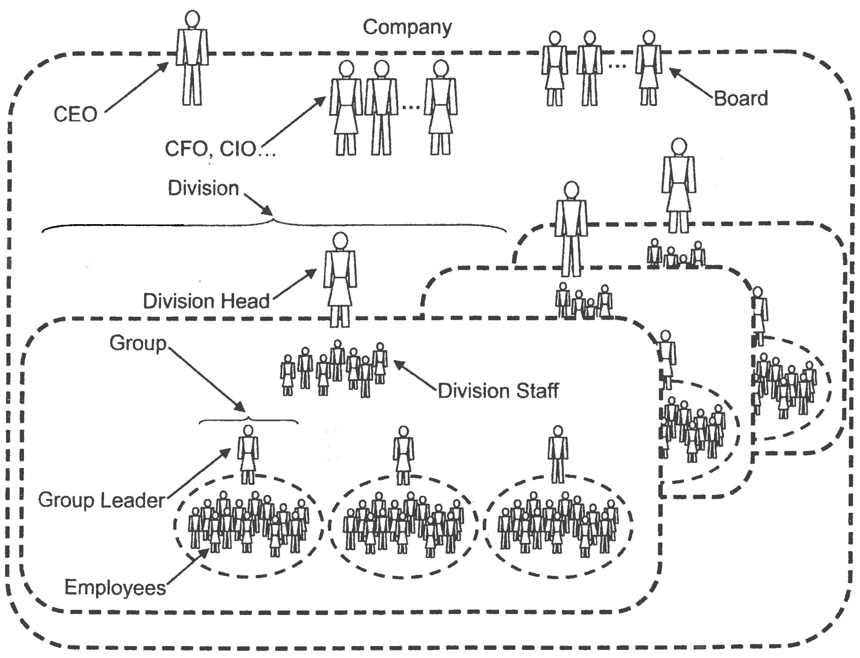
1. A finite undirected graph G(V,E) with n nodes (or vertices), representing objects or agents, and m edges. G also has no multiedges or self loops.
2. A domain D of state values, representing the properties of each node.
3. A local transition function set $F=\{f_1, f_2, \ldots, f_n\}$, with each f_i mapping D^{δ_i+1} into D, where δ_i is the degree of node i. (f_i takes each node and its neighbors at a state, and yields a new state for the node).
4. A total order π on the set of nodes, specifying the order in which they update their states.

A configuration K of the system is an n-vector $(k_1, k_2, \ldots, k_n)$, where each k_i is an element of D. In a time-varying sequential dynamical system, the topology of the graph or the local functions can vary or evolve over time, as well as the configuration of the sequential dynamical system.[8]

Even the simplest models, as I mentioned, involve heterogeneous sets as agents. This is a significant advantage over analytic models that are limited to a single representative agent or a very small amount of heterogeneity if they are to be tractable. More generally in agent-based modeling, objects are not just simple things but have a variety of properties, behaviors, and tasks they can execute. The properties of agents are usually intrinsic properties, and the behaviors are either internally driven or can be triggered by changes in the objects with which they are connected. Connections between agents are causal: one agent does something, or has a property, and that triggers a property change in another agent. Moreover, environmental factors are also just more agents, so even in simple models agents representing individual people can interact dynamically and bidirectionally with agents representing environmental properties as well as with one another.

In some sense, agent-based models can model objects at multiple scales. The widely used modeling program Swarm, for instance, allows agents to be aggregated sets of other agents. Likewise, a prominent recent book on agent-based modeling in business depicts organizations like companies as being built up out of divisions, which are built out of groups, which are built out of individual employees.[9]

Given such a compositional hierarchy, typical agent-based models take to the behavior of the high-level agents to be determined from the "bottom up," with the behaviors of the high-level agents determined by the behaviors of the components making them up. In Swarm, for instance, one might model a population of rabbits, each composed of rabbit parts, each of which is composed of cells. When a rabbit receives a message from the scheduler governing that agent, the behavior of that agent is determined by



Multiscale organization with employee subagents.

Figure 6.3 Multiscale organization with employee subagents. Courtesy of Oxford University Press from *Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation* by Michael North and Chares Macal (2007).

the behaviors of the swarm of the agent's constituent agents.[10] Inasmuch as an agent is constituted by a set of lower-level agents, the "causal efficacy" of the high-level agent is fully exhausted by the causal efficacy of the constituent agents.[11]

The successes in agent-based modeling, in such areas as epidemiology, power markets, and combat simulation, may lead us to be confident that we will be able to extend agent-based modeling more generally. But as soon as we move from these areas to more typical social properties, I will argue, the methodology becomes less effective. Related shortcomings apply just as much to analytic models as they do to agent-based models. But it may be mistakenly thought that agent-based modeling is immune to these problems, and the problems can be raised particularly clearly in the context of agent-based models.

TWO FALLACIES OF INDIVIDUALISM

"Ontological individualism" is typically taken as a truism in the philosophy of social science and is a background assumption of both analytical and computational models in the social sciences. Ontological individualism is a metaphysical claim about the relation of social facts or properties to individualistic facts or properties.[12] It is usually cashed out in terms of a form of the supervenience relation, to analyze the "local" claim *the social properties of any entity exhaustively depend on that entity's individualistic properties,* or the "global" claim *the social properties holding in a world depend on the individualistic properties holding of and among individual people in that world.*[13] However, ontological individualism—even understood as a global supervenience claim—does not just fail to be a truism but is false.[14]

The two fallacies I will discuss below are both fallacies about the ontology of social properties. They are what we might call "anthropocentric" fallacies, arising as a result of overestimating the extent to which social properties depend on individual people.[15] Obviously social properties are introduced by people. That does not mean, however, that these properties are fully determined by (or are exhausted by or supervene on) individualistic properties. The sources of this falsehood can be traced to two different problems, as I will discuss just below.[16] Before introducing them, let me begin with a brief comment on the relevance of ontological considerations to modeling.

Ontology in Modeling

While it may not be obvious that ontology is pertinent to modeling at all, ontological assumptions at least are tacitly built into the construction of any model. Suppose one wants to model a pot of water boiling on the stove.

A natural first step is to break the system into parts—e.g., the water, the pot, and the burner. One might then model the water with a molecular dynamics simulation, the pot with some analytic heat equations, and the burner with a Monte Carlo simulation.

Different models of some entity may treat the same entity in incompatible ways, such as modeling water as an incompressible fluid when modeling hydraulics and as a collection of particles when modeling diffusion.[17] But in a single model of physical stuff, we typically want the constituents to be mutually exclusive or nonoverlapping. In constructing a model of water boiling on a stove, the modeler thus breaks it down into parts that interact with one another, avoiding redundancy or internal conflict in the model. Once the modeler has constructed the analytic heat equations to treat the pot, she does not want to redundantly include the pot molecules in the molecular dynamics simulation that is treating the water molecules. It is usually straightforward to accomplish that.[18] All the ontological work the modeler does is to draw a spatiotemporal boundary around each object and then treat their causal interactions with one another. However, as I will discuss in a moment, typical social properties are different from typical physical properties in this respect.

To discern the role of ontology in model construction, it is important to retain a clear distinction between ontological dependence and causal dependence. The temperature of the water is causally affected by (and we might say "causally depends on") the temperature of the pot. But the water does not depend ontologically on pot molecules, whereas it does on the water molecules. Changes in the water consist in nothing more than changes in the water molecules; they are not caused by them.[19]

The Locality Fallacy

Many physical properties depend (ontologically) on spatiotemporally local features of the objects they apply to. Whether I am hot or cold depends on the temperature of the molecules of which I am materially constituted.[20] Some social properties are also local to me. My having the property *dancing an Irish jig* depends only properties spatiotemporally local to me. I need to be moving my body in a certain way, and there needs to be a floor beneath me which I am tapping in a certain way. Although the conditions for what it takes to have the property *dancing an Irish jig* were defined by certain Irishmen, the conditions for my having the property only involve my body and a small region around it.

A great many social properties, however, are not locally dependent. This point has been noted by, among others, Currie 1984, Ruben 1985, and Pettit 2003, though they did not draw out the implications for social theorizing or modeling. An example of a property that is not locally dependent, for instance, is the property *being President of the United States*. The fact that Barack Obama has that property does not depend on his intrinsic properties, nor does it depend on the properties of the White House, the places he has traveled, or the people he has come in contact with. Rather, it depends on a variety of properties of other people and things. For instance, it depends on certain current and historical properties of the electoral college, and the facts on which those properties depend in turn, as well as on such facts as the existence of the United States. and the U.S. government, etc., and all the factors on which those depend in turn. For modeling purposes, many of these can often be left in the background, as a practical matter. But nonlocal dependence may in many cases be important for model construction.

The "locality fallacy" is the fallacy of taking a nonlocal property to be a local property, i.e., taking some property P holding of an object x to depend (ontologically) on factors local to x, when P's holding can in fact depend on factors that are not local to x. A model can implicitly commit the locality fallacy if in modeling a nonlocal property, the only factors it takes into account are local factors and causes that impinge on the local factors. To be sure, many social properties that are in fact nonlocal can be usefully modeled, for many purposes, as if they are local properties. Similarly, it is often possible to do a nice job modeling a pot of water boiling on the burner by modeling only the burner and the water, while overlooking the pot. At the same time, if *all* models of the pot of boiling water, even very detailed ones, completely ignore the pot, we might ask whether the pot is being overlooked because we have a blind spot about pots. Likewise, local models of properties that are ontologically dependent on nonlocal factors are fine for many purposes. But if our approach to modeling both systematically and unconsciously ignores nonlocal factors, it is reasonably to ask whether this is a design flaw in our approach.

Many social properties are nonlocally dependent in very straightforward ways. Consider, for instance, an obviously extrinsic property of a social group, such as being in the National League playoffs, or being charged as a corrupt organization under the Racketeer Influenced and Corrupt Organizations (RICO) act. The factors determining the holding of such a property depends on factors that are not local to the bearer. But for such cases, it is unlikely that a modeler would fall prey to the locality fallacy. The troublesome cases are the ones where the nonlocality is not so straightforward.

Consider, for instance, the fact *the average age of the Tufts freshman class is 18½*. This obviously depends in part on certain local properties of a collection of approximately 1300 individuals. But only in part: that fact is not local to those 1300 individuals, for similar reasons that *being President* is not local to Obama. To be a member of the freshman class depends on many factors that the freshmen may not even be aware of or that may not even impinge on them causally. Suppose, to simplify, that there are only four freshman, P, Q, R, and S, aged 17, 18, 19, and 20, respectively. Evaluating *the average age of the freshman class* in the actual world, we consider the ages of P, Q, R, and S, which average to 18½. And suppose that all four of them go to a day-long lecture one day during the fall term. Over the course of the day, their individualistic properties, including their ages, remain relatively unchanged. But imagine that while they are sitting in the auditorium, the world changes radically around them: at 10 a.m., P's parents and Q's parents win the lottery and immediately withdraw their kids from school; at 11 a.m., S's parents go bankrupt and withdraw S from school as well; at 1 p.m., P's parents have second thoughts and re-enroll P; and then at 3 p.m., the board of trustees dissolves the school entirely. Over the course of the day, the individualistic properties of P, Q, R, and S remain more or less constant, but the value of the function *the average age of the freshman class* fluctuates, from 18½ at 9 a.m., jumping to 19½ at 10 a.m., dropping to 19 at 11a.m., dropping further to 18 at 1 p.m., and then becomes undefined at 3 p.m. This function fluctuates in virtue of changes in properties other than the individualistic ones of the freshmen themselves. It is not that the values of these functions do not depend on the properties of the four individuals, but rather, that they also depend on those nonlocal properties that figure into determining the holding of the property *being a freshman*.

Not every social property of an individual or group is nonlocally dependent. Consider the choices of a pair of prisoners, each given certain information and certain alternatives. Then the only factors on which the output of the "choice" function applied to the pair of prisoners depends are their local characteristics. The same is true for the audience in an auditorium in the example Thomas Schelling discusses in the introduction to Schelling 1978. To determine why an audience has spontaneously organized to sit bunched together in the seats at the back of the auditorium, as opposed to populating the better seats, the only factors that pattern depends on are again the local characteristics of the individuals in that audience. The reason is that the property *being in an auditorium*, like *being a molecule in a balloon* and unlike *being a freshman*, plausibly depends only on the characteristics of that local spatial region.

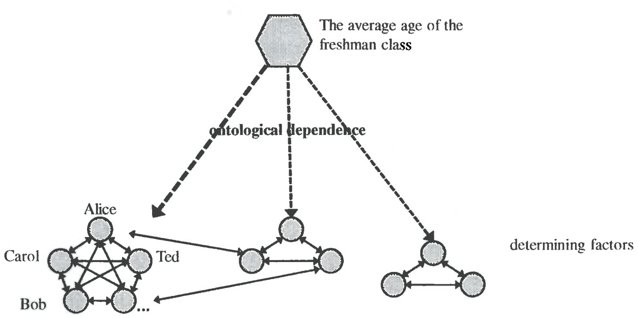


Figure 6.4 Ontological dependence and causal interactions.

In Figure 6.4, the dotted arrows represent lines of ontological dependence, and the thin solid arrows represent causal interactions. The value of the function *the average age of the freshman class* ontologically depends on a variety of individualistic properties, some of the bearer of the property, represented as the pentagon on the lower left, but also depends on properties of other members of the population apart from the bearer of the property. Some of these may causally interact with the members of the class, and some may not. Given the ontological dependence of the macro-property on the wider population, it is possible to change that macro-property through a change in a nonlocal property, even if that nonlocal property does not even causally interact with the freshmen who are the bearers of the property.

One might still have reservations about the pertinence of nonlocality to modeling social properties such as *being President*, because *being President* seems to have the following characteristic. Although it is put in place by a variety of nonlocal factors, such as being elected by electors, once that has taken place, those nonlocal factors seem to become irrelevant. Although the choices of the electors were relevant in making Obama President, he remains President even if the electors or the population as a whole now have changed their attitudes. This suggests that subsequent to the election, things like *the decisions of the President* or *the actions taken by the President* depend only on his local properties, at least until the next election.[21]

It is true that a change in the attitudes of voters across the country does not suffice to discharge Obama. But that does not mean that the nonlocal factors on which *being President* depends are irrelevant to models of the property. Even after he has been elected the determining factors may change. It is not only having been elected, that Obama's having the property *being President* depends on, other factors, such as not having been impeached and convicted, having the government in place, not having been removed in a coup, and so on. Many changes external to Obama could make the property *being President* fail to hold of him, whatever his local characteristics. It may be acceptable to relegate these to the background in a short-duration model of the U.S. presidency, but doing so may be a poor strategy for modeling different countries or time periods or models of longer durations. For other extrinsic properties, it may be even less satisfactory to leave all nonlocal factors in the background.

As with all modeling, one always must be selective about the factors to include. For certain models of *being President* it may be useful to ignore the nonlocal factors on which the property depends, just as we might ignore nonlocal factors in modeling the average age of the freshman class. But for different purposes, it may be preferable to ignore the local ones and model only the nonlocal ones. To return to the boiling water analogy: for some models it is wise to ignore the pot. But for other models, it may be the burner, or even the water, that is practical to neglect, in favor of the pot. It is not clear why there should be a systematic bias for incorporating the local factors on which social properties ontologically depend and for ignoring the nonlocal ones.

The Levels Fallacy

The levels fallacy is an equally pervasive problem in models of social properties. Roughly, it is that if we want to model entities or properties at the social level, we only need to decompose them into entities or properties at the level of individuals. This seems to stem from the same sort of reasoning that motivated the Oppenheim-Putnam picture of the sciences being divided into compositional "levels."[22] Their idea was that objects can be arranged in a compositional hierarchy of levels: (i) elementary particles, (ii) atoms, (iii) molecules, (iv) cells, (v) multicellular living things, and (vi) social groups. The sciences correspond to the study of objects at these levels. Among the conditions they imposed on a hierarchy of levels were (i) for something to be in a particular level of the sciences (except the lowest level), it needs to be fully decomposable into objects of the next lower level; i.e., that each level is a "common denominator" for the level immediately above it, and (ii) nothing on any level should have a part on any higher level.[23]

There are a number of problems with this picture. First, it is clearly not the case that objects at a given level do fully decompose into objects at the next lower level. The picture is reminiscent of the belief of early cell biologists[24] that the human body was exhaustively composed of cells. That, of course, is not the case: we are made up partly of cells, but we also are made of digestive fluids, blood plasma, bone matrix, cerebrospinal fluid, intracellular material, mucus, and so on. Bodies do not decompose into objects only at the cellular level.

Likewise, it is a common but glaring mistake to think that social objects decompose into individual people. Consider, for instance, Tartine Bakery on the corner of 18th and Guerrero in San Francisco. On a typical day, racks of croissants are being baked in the ovens, the cash register is ringing, bakers are working with flour and sugar and butter, customers are lining up out the door, credit cards are being processed, banks are being debited and credited, accountants are tallying up expenses, ownership stakes are rising in value, and so on.

The employees are critical to the operation of Tartine. Is it, however, plausible that the employees exhaust the parts of the bakery? The employees are plausibly part of the bakery: if there were no bakers, then it would arguably be an empty shell and not a bakery. But a bakery also needs ovens, and the ovens are parts of the bakery. It is further plausible that among the parts are also its butter and its flour, and even its cash registers and coins and bills exchanging hands. Historically, there have been various attempts to force-fit the dependence-base of social entities and properties into some preferred class, supposedly in the service of individualism. Most notably, behaviorists argued that social entities are exhaustively composed of the nonintentional behaviors of individual people, and psychologistic approaches took social entities to be exhaustively composed of psychological states. These, mercifully, have largely died off, but the assumption of an individualistic base seems to linger on. The neat hierarchical composition of corporations out of nested groups of individuals shown in Figure 6.3 is an example. That picture overlooks everything but the employees.

Equally significantly, even if the compositional pyramid were correct, that would not mean that social facts depend exhaustively on individualistic facts. Simply because a group is constituted by individuals does not imply that the group is identical to its individuals. Work on "coincidence" in recent years has argued for a gap between the constitution and the identity of ordinary objects, so that two entities may be identically constituted, and yet have different actual and modal properties.[25]

Consider the factors on which some social property, such as *having a $100 million liability for hurricane damages in New Orleans after Hurricane Katrina,* depends. A large number of factors determine whether or not such a property holds of some insurance company.[26] Among the factors may be the psychological states of employees of the company, but another factor their liability depends on is the actual hurricane damage in New Orleans. Corporations, universities, churches, governments, and so on, depend on a wide variety of objects, including contracts, liabilities, insurance policies, stocks, bonds, etc. These objects, in turn, plausibly depend on complex sets of other objects. Even the stratification of objects into levels itself is dubious: at best, the division of objects into levels is fuzzy and pragmatic.[27]

Instead, social entities and properties depend on a radically heterogeneous set of other entities and properties. They can be cross-level, they can be at multiple scales, and they do not have to be individualistic at all, as depicted in Figure 6.5.

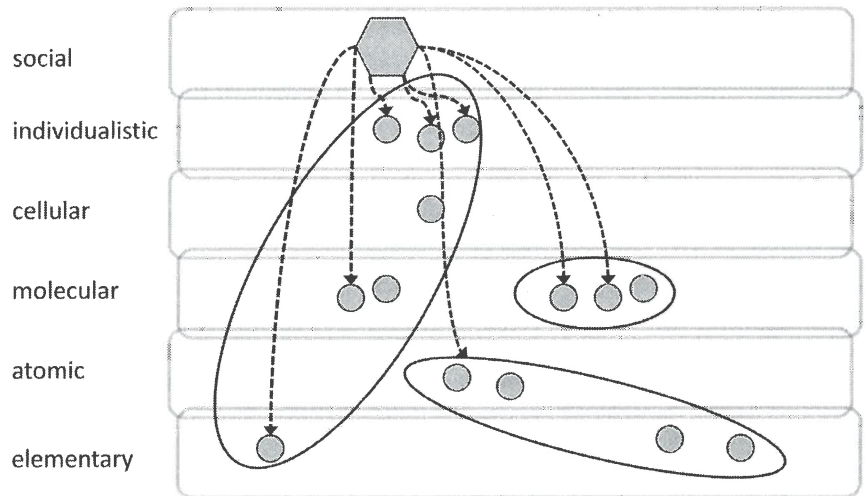


Figure 6.5 Radically heterogeneous set of other entities and properties.

In the figure, a social property such as *having a $100 MM liability for Katrina damage* may depend on certain properties of the insurance company bearing the property, represented by the large oval in the diagram. And it may depend on other objects, such as buildings, swimming pools, cats, vats of bacteria producing biofuels, particle accelerators, and on and on, represented by other ovals in the diagram. Again, these dotted arrows are arrows of ontological dependence, not arrows of causation.

AGENT-BASED MODELING AND THE FALLACIES

The locality and levels fallacies are mistakes pertaining to the relation between macroscopic facts and microscopic facts. For some facts, however, these issues may not arise. Let us call macroscopic facts *simply dependent* if they are local and depend solely on the next lower level of facts.[28] How much sense this makes depends, of course, on how much sense we can make of levels in the first place. But as we proceed, it is useful to contrast models in which it is assumed that all macro-properties to be modeled are simply dependent, from those that are not.

Agent-based models have the resources to avoid the naïve fallacies of individualism, by incorporating heterogeneous ontologies of agents. But the nonlocal and cross-level dependence of social objects and properties is often overlooked entirely, so they do not even avoid the fallacies in naïve forms. More significantly, however, there are limits to how much even sophisticated models can avoid the fallacies.

In the following sections, I argue that the problems arising from the locality and levels fallacies create a set of new problems for agent-based modeling, beyond those that arise for modeling simply dependent facts. After setting up some distinctions and terminology in the next section, in the following section I consider some issues that occasionally arise in modeling simply dependent facts and can usually be addressed though judicious choice of variables to model. I then argue in that nonlocal and cross-level dependence turn these minor issues into apparently insurmountable obstacles. Subsequently I turn to refinements of agent-based modeling, proposing ways of designing models for addressing these problems but at the same time having to make certain new compromises.

Distinguishing the Target, Base, and Causal Ontologies

In the design of a model, the modeler implicitly has a set of entities of interest, or "target" entities, in mind, although it may not be apparent from the implementation. An agent-based model targeting a macroscopic entity or property M will generally result in different choices than one targeting some other entity or property M'. Given a set of target entities, the first step in model design is to identify some set of microentities on which that target set ontologically depends. These factors we will call the "base ontology."[29] The next step, then, is to identify the other entities that causally interact with the base. These factors we will call the "causal ontology."

For a given target, we sometimes model just a fragment of the base ontology, and often (if not always) model just a fragment of the causal ontology. Given a target ontology T, I will distinguish the "modeled base ontology" M^B from the complete base ontology of T and distinguish the "modeled causal ontology" M^C from either the complete set of factors interacting with the complete base ontology, or the complete set of factors interacting with the modeled base ontology. Together, I will call $M=M^B \cup M^C$ the "modeled ontology."[30]

If we suppose that the macroentities being modeled are simply dependent, the base ontology can be entirely local to the target, and consist of entities at level n-1.[31] For a simply dependent target, the modeled base ontology can thus often be the entire base ontology. Given the base ontology, the "modeled causal ontology" is chosen on the basis of being entities in the world that are expected to interact in a significant way with those in the base ontology.

Only when the target is simply dependent is the relation between it and the modeled base ontology so straightforward. When the target is not simply dependent, the determination of the modeled base and modeled causal ontologies from the target becomes far more problematic.

The Dynamics of Ontology Under Simple Dependence

Even with the assumption that the target is simply dependent, there are some interesting problems arising from the fact that the relevant causal factors and even the relevant dependence base may change over the course of a model's evolution.

In a standard graph-dynamical system, like a sequential dynamical system, the configuration K (i.e., the state of the nodes) changes over time, but the graph itself is static. A time-varying sequential dynamical system accommodates the fact that the interactions among objects are not static, as the system evolves and across circumstances. For both, however, the modeled ontology itself is fixed at the outset.[32]

Where it is not known in advance how much causal influence some factor will have, a natural approach is simply to include it in the causal ontology. If the modeled ontology is fixed at the outset, we simply have to choose a big ontology even if we know that an object being modeled has a predictably changing causal ontology over time. For instance, to model a car traveling down a highway, where the things the car interacts with at time t are different from those it interacts with at t', a natural if inefficient treatment is to expand the ontology to include the larger set of causal features the system will have contact with over the duration of the simulation.

For systems that are well-behaved,[33] inefficiencies of fixing a causal ontology in advance are often not overwhelming. But even for simply dependent properties, changes as the system evolves may mean that new factors in unpredicted locations or at unpredicted scales may become relevant. As a system develops, the fragment of the causal ontology that is included in the model may thus need to change, in order for the model to be both reasonably tractable and reasonably accurate. In other words, we may not only need the network of interactions to be time-dynamic but the nodes as well. Rather than just a single modeled causal ontology M^C, the causal ontology may need to become a sequence (M^C_0, M^C_1, . . .).

Although this seems simple—just modify the representations of time-dynamic graphs to allow the nodes to change over time—it creates a potentially severe problem. Consider a modeler observing her model during a run, evolving over time from steps $t_{0 \ldots} t_{i \ldots} t_n$. Suppose that at time t_i, the modeler realizes that the model has evolved in such a way that some entity that was not included in $M^C_0 \cup M^C_1 \cup \ldots \cup M^C_{i-1}$ will have significant causal influence. For instance, suppose she is modeling the orbits of planets and finds that one planet is thrown out of its orbit partway through a run and enters a different system that she had assumed she could neglect. Accommodating the new elements into the causal ontology may not be as simple as just adding them into M^C_i and letting the run continue. Because, of course, over the simulation up to t_i, it might have been influenced by other entities in the domain, and given those influences it will have reciprocal effects on the other parts of the ontology, given the way the system in fact evolved. To model it properly, the run may need to be restarted from scratch, to incorporate the missing node and its influences. Even admitting small changes to the causal ontology may turn a computationally tractable model into one that must be repeatedly rolled back.[34]

Furthermore, even with a simply dependent target, it may not only be the causal ontology that needs to be dynamic, but the base ontology as well. If we are modeling the gas escaping from a chamber or the traffic at an intersection, for instance, the composition of the target changes over time. This means that the base ontology is also not static but perhaps should become a sequence (M^B_0, M^B_1, . . .).

For simply dependent targets, however, these issues tend to be easily accommodated. In many such cases, a changing base ontology does not require a change in the modeled ontology at all, since the changes in the base ontology simply involve a shift from a factor being part of the modeled causal ontology to being part of the modeled base ontology. That is, the overall ontology does not need to change if the modeled ontology M already includes the factors left out of M^B_t. This occurs if the difference between M^B_{t+1} and M^B_t is a subset of M^C_t. One of the reasons that modelers do not make a distinction between the base and the causes is that it can actually be an advantage to conflate the two, when factors can flip in and

out of the base ontology over time. But as we will see, failing to distinguish the base ontology from the causal ontology is acceptable only so long as the base ontology itself is well-behaved.

The reason these considerations do not foreclose the possibility of modeling altogether is that we generally choose the entities to model judiciously, and for these entities, the world is reasonably well behaved. Simply dependent target facts do not guarantee good behavior, but inasmuch as we are dealing with systems that are basically dependent on local and single-level factors, the dynamics of causal ontology are often sufficiently predictable.

With many social properties, however, these assumptions should not be expected to hold at all. When the target is simply dependent, the dynamics of the base ontology, and even the dynamics of the causal ontology are typically small issues. But they can balloon in importance when the target is not simply dependent.

When Targets are not Simply Dependent

If the target is not simply dependent, it may be a problem to choose a reasonable modeled ontology altogether. Intuitively, the problem with nonlocal dependence is that the dependence base can become indefinitely *broad*, and the problem with cross-level dependence is that it can become indefinitely *deep*.

A non-simply dependent target may depend on factors that are spatiotemporally remote from it. To model a property such as *being President of the United States*, for instance, facts about the entire voting population may need to be included. These factors may change over time and across circumstances, meaning that the modeled causal ontology cannot be restricted to factors that are spatiotemporally local or connected to the President.

Even more significant, however, is the implication of non-simple dependence on the choice of a modeled base ontology. With changes over time and across circumstances, the choice of the modeled base ontology can have to change radically and discontinuously, if it is to be a reasonable dependence base for the target. Even for the simple property *being President of the United States*, this may be the case. This property depends on certain properties of the voting population. Yet the voting population is not static. *Being President of the United States*, for instance currently depends in part on the votes of a large group of 18- and 19-year-olds who were not in the electorate a few years ago. The dependence base is significantly different today than it was then.

If there is single-level dependence, this may not be disastrous. Recognizing that social properties such as *being President of the United States* may depend on all the members of the population, we may simply choose to model the entire population. There are not so many people in the population, that the problem becomes computationally intractable. The failure of single-level dependence, however, erases this possibility.

For example, consider some complex but important social fact that we may want to model, such as the fact *American International Group (AIG) has $2.3 trillion in credit default swap obligations*. This ontologically depends on a great variety of factors, including properties of houses, paper, contracts, other corporations, its employees, assets backing various bonds, etc. Suppose, for instance, that a third of that obligation involves being contracted to replace the payments from mortgage-backed bonds, should the issuers default. Simplifying a bit, each of the mortgages backing those bonds consists of a disjunction—either a stream of payments from the homeowner or a house that the issuer has the right to seize if the payments stop. American International Group's instantiating the property *having $2.3 trillion in credit default swap obligations* thus depends in part on the houses. To model this as an intrinsic or psychological property would grossly mischaracterize the factors determining whether the property is instantiated. Instead, the factors on which the property ontologically depends can be any of thousands of different aspects of the housing stock, other banks, and so on. All of these are eligible to be components of the modeled base ontology.

It is impossible to model the entire world in infinite detail, so the modeled base has to be a judiciously selected subset of the complete base ontology. As I noted above, if we are modeling a simply dependent property, such as the density of traffic at an intersection, then changes to the base ontology are usually easily accommodated. But when the target is not simply dependent and instead depends on highly nonlocal and cross-level factors, the choice of a modeled base becomes far more problematic. Moreover, it becomes increasingly likely that as the model runs a much more varied set of factors will become relevant. The factors that a modeler might reasonably incorporate are likely to change nonlocally and are also likely to change at different depths of texture: whereas at time t the modeler may have reasonably ignored everything at the level of housing stock, water supplies, or banking charters; any one of those might be the most relevant aspect of the base at time t+1. This problem does not arise for simply dependent properties.

Still worse, when these changes in the modeled base ontology run into the choice of a causal ontology interacting with the base, the problem explodes. Normally, for each M^B_i a corresponding causal ontology M^C_i would to be chosen, incorporating the factors that significantly causally influence the entities in M^B_i. But if M^B_i is spatially and level-wise ill-behaved, this makes the sequence of significant causal influences at least as badly behaved. And further, it increases the risk of the need for rollbacks. If even minor additions to the causal ontology potentially required rollbacks, with the present problems we might never be able to get past the first few steps of a simulation, without having to restart over and over again.

The Hard Problem: How Even to Determine the Base Ontology

These threats to the tractability of a social model are potentially game-enders. But there is a more difficult problem still. Given a total ontology M_t consisting of a base ontology M^B_t and a causal ontology M^C_t: what is M_{t+1}? The problem is that presuming that M^B_t is chosen to be a reasonable subset of the complete dependence base of the target ontology for use at time t, there is no reason to suppose that the state of M^B_t (or M^C_t) has enough information to determine or even recommend what a reasonable M^B_{t+1} might be.

Under the assumption that the target is simply dependent, the modeler could be generous about the factors at a single level to be included in M_t and confidently predict any that changes in the modeled base and modeled causal ontologies would be local or nearly local to M_t. But without simple dependence, the realization of the target in the dependence base can change any which way, both broadly and deeply. A choice of the dependence base to model at a given time, as a subset of the complete dependence base, does not determine which subset of the (potentially global) dependence base of the target should be chosen subsequently.

To see this, consider again the AIG example. Which factors in the complete base ontology are significant to determining the target can differ quite radically even in relatively nearby worlds and times. And the nearest set of circumstances to the actual one might be circumstances that involve a radically different set of nonlocal determining factors. It may be the following are three likely paths the world could take, that would potentially change the fact *AIG has $2.3 trillion in credit default swap (CDS) obligations*:

1. Houses fall into disrepair because homeowners do not have the incentive to keep them up, and hence lose value at a rapid rate, and so on . . .
2. Depositors are losing confidence in banks and withdrawing their money en masse, causing them to collapse, . . .
3. A rash of resignations among executives are leading the AIG board to choose to default on various instruments, . . .

These are three very different possible paths that the world might go in, which may change the fact in question. Note that these do not only involve causing changed states of the entities in the dependence base but may involve radically different dependence-bases themselves, as circumstances change. In path 1, houses which were once constitutive in part of the obligation may be destroyed; in path 2, the houses may remain but the counterparties to the transactions may disappear; and in path 3, contracts that were in part constitutive of the obligation may be torn up.

From a distance, we can see changes in the base coming: the housing stock, the banks, and the stress among executives are being modeled, each of which may flag a change in the modeled base. But while those factors are externally observable by modelers with the target ontology in mind, the model itself at t only includes the modeled ontology. It does not include the elements of the complete base that will be significant factors in determining the target at t+1. Nor do the factors in the modeled base determine which other factors will be significant parts of the dependence base of the target at t+1. The dependence base of the fact *AIG has $2.3 trillion in CDS obligations* at t+1 is not necessarily included in M_t, nor is it determined by M_t, regardless of how judiciously M_t is chosen.

Ignoring or minimizing changes in the base ontology may be a reasonable idealization for simply dependent targets. But when we move even to straightforward properties that are not simply dependent, these idealizations can collapse spectacularly.

REFINING AGENT-BASED MODELING

At this point, things seem bleak. The modeler does her best to choose a starting point for the simulation, based on the target entities of interest. But the ontology needs to be dynamic, in order to accommodate the changes in the realization of the targets, and every dynamic change in the ontology threatens to demand a rollback of the model to the beginning. On top of that, the modeler herself, in order to determine how the base ontology changes over time, needs to keep appealing back to the target. She can have as finely grained a model as she likes, and still the target entities cannot just be put to the side, to be read off of the model, as it evolves. It seems the only routes are either to accept the approximation of all targets as well-behaved simply dependent entities—e.g., treating AIG's liability as if it depends on features of the AIG employees alone, which may be tantamount to modeling a Ferrari by simulating its air conditioner—or else to model everything at all levels in infinite detail. If so, we might just throw up our hands.

This pessimism, I think, is unwarranted. Instead, I suggest that we give up on two deeply held assumptions I mentioned at the outset: first, that agent-based (and other) models provide the micro-foundations for macro-properties and second, that models have to avoid ontological redundancy. Abandoning each of these is painful but may be less costly than the alternative.

Proposal: Explicitly Including the Macroscopic Properties in the Ontology

Across circumstances and time, the relevant dependence base of a target may be heterogeneous and volatile. If we do not model the complete dependence base of the target (i.e., much of the entire world in infinite detail),

then the determination of the dynamics of the base ontology requires appeal to the macroscopic target itself. But it is possible to include the target and other macro-properties in a model, so long as we are willing to make some sacrifices.

Consider *AIG having $2.3 trillion in CDS obligations,* given a particular modeled base ontology M^B_i for it at time t, and a causal ontology M^C_t. To determine M^B_{t+1}, we need to consider the dependence base of the target property given the state of M^B_t. For instance, if the system has evolved such that the banks are nearing default at t, then M^B_{t+1} will be chosen to involve the factors determining bank defaults. This can be done computationally—but just not on the basis of M^B_t alone. Instead, the target entity can be included in the model, not just "read off" of it, as an object or agent with causal efficacy and causal factors affecting it.[35]

In a sense, this proposal is a radical capitulation. At the outset, I mentioned that the basic idea of an agent-based model has seemed to many people to be the provision of micro-foundations for macro-properties. The explicit inclusion of macroentities in the model undercuts this aspiration. But in light of the nonlocal and cross-level dependence of social entities, I suggest that true micro-foundations—even computational and nonreductive ones—are a pipe dream. There is no reason to think that most social properties are well behaved with respect to a fixed or tractably dynamic set of microentities on which they depend. And since a model of every detail of the entire world is possible only in science fiction, the exclusion of macro-properties from models entails a reduction in their accuracy.[36]

Surreptitiously, this may already be taking place in agent-based modeling. It is likely that many agent-based models implicitly include macroscopic factors in the models, even while they seem to be micro-foundational. For instance, in models of traffic patterns, parameter setting is frequently done by tweaking them until they generate the desired macroscopic properties. As far as I know, no one has investigated the question whether this is illicit when applied to agent-based models and whether it compromises micro-foundations. But there are many ways macro-properties can be hidden under the covers in a model, and it may not turn out to be such a bad thing.[37]

The Problem of Redundancy

Including macro-properties in models, however, carries with it a different and rather significant risk: the potential for redundancy in the model. In the model of the boiling water, the modeler implicitly makes sure that her base ontology is nonredundant. If she decomposes the water into individual molecules, those supplant entities that are already composed of individual molecules, such as waves and eddies and vortices.[38] If both macroelements and the microelements composing them are included, both with causal effects, there is a risk of double-counting causes. If both *the CDS obligations of*

AIG and *the housing stock on which it partly depends* are included in the modeled ontology, there are subtler but similar risks of redundancy.

This is a large problem, but not necessarily an overriding one. When properties are simply dependent, there may be no reason to take the risk of double-causation. But the exclusion of macro-properties from a model, when the base ontology is not well behaved and when we do not model the entire world, threatens how accurate the model can be anyway.

Agent-based models have always traded off computational tractability with the complete inclusion of potentially relevant factors in the modeled ontology. They have not thought to compromise nonredundancy. But nonredundancy is simply another desideratum, the partial sacrifice of which does not spell the complete failure of a model. If the threats of redundancy can be kept under control, then it may be worth it to admit some redundancy in the model so as to improve the choice of factors to be modeled over time and across circumstances, while retaining computational tractability.

Doing so also requires keeping track of the target, base, and causal ontologies separately, with the constitutive graph and the causal graphs separately represented, rather than collapsing all the nodes into a single modeled ontology, and taking all the graph edges to be causal interactions.

MODERATING THE PICTURE OF AGENT-BASED MODELING

As I mentioned at the outset, agent-based models are advantaged in many ways over analytic models, in terms of how comprehensively and realistically they can model the world. It is typical for agent-based models to incorporate a variety of environmental factors, even factors that are not necessarily local to the people in the population. The objects taken as the components of the model can be much more general than the individualistic factors employed in a typical model in analytical economics. Agent-based modeling is not strictly individualistic, so it does not have the same shortcomings of individualism. Sometimes, individualism is presumed in implementations of agent-based modeling, since individuals are regarded as the locus of social properties, and there is a failure to distinguish constitution of an entity from the factors that determine its properties. But this is not an in-built limitation of agent-based models.

Agent-based models have well-known disadvantages as well, of course, such as the difficulty of interpreting, testing, and drawing generalizations from them. The problems I have discussed here, however, are equally applicable to analytic models as they are to agent-based models. It is because agent-based modeling does seem to have an ontological advantage that it is most useful to point out the ontological problems with social modeling in this context. Inasmuch as correcting for the fallacies of individualism means changing how agent-based models are built, *a fortiori* it means rethinking analytical models as well.

To summarize some of the implications for agent-based modeling: First is the value of making the choice of target ontology explicit. The representations of the target, base, and causal ontologies should be kept separate, as should the representations of the dependence graph and the causal graph, given that all can evolve separately. Given an explicit target, the complete base ontology ought to be understood, so that a modeled base ontology can be judiciously chosen and changed as the system evolves. The base ontology, of course, can include heterogeneous properties and agents, nonlocal and cross-level. To accommodate a changing base ontology, macro-properties may need to be included explicitly in the model, not just as rollups of microagents, but having causal interactions. In modeling changes in the ontologies, as well as the causal evolution of the states of the system, some pre-modeling may be employed and/or allowances for rollback made, as the evolution of the system forces reassessment of the most significant dependent and causal factors. Finally, nonredundancy may be traded off explicitly against the goal of ensuring that the modeled ontology is well chosen over time, in idealizing the model so as to maintain its computational tractability.

APPENDIX: ON THE FAILURE OF LOCAL SUPERVENIENCE

In the paper, I describe the "locality fallacy" as the fallacy of taking a nonlocal property to be a local property. The reason this is a fallacy is that many social properties in fact ontologically depend on "nonlocal" factors. In this appendix, I discuss this dependence in more detail than above, and in particular, consider how "local supervenience" typically fails for social properties.

To begin, it is crucial to note that "local" is being used in two different ways in the present discussion. In the body of the paper, I used "local" to characterize a certain kind of property of an object. As I mention in Note 20, I mean the idea of a property being "local" to be somewhat looser than a property being "intrinsic." Intuitively, an intrinsic property of an object is a property that holds or fails to hold regardless of what other objects there are in the world. A local property of an object is one that holds or fails to hold regardless of what the world is like at a spatiotemporal distance from the object.

In speaking of supervenience claims, the term "local" is generally used differently. Supervenience relations are typically defined as holding between sets of properties—a set of "A-properties," such as mental properties or chemical properties, supervening on a set of "B-properties," such as neural properties or physical properties. In *local* supervenience claims, the dependence of A-properties on B-properties is cashed out as a comparison between pairs of objects in any possible worlds. Taking any possible pair of objects, a difference in the A-properties of the pair implies a difference in the pair's B-properties. Notice that a local supervenience claim might be true even if it is not intuitively "local" at all. For instance, it is trivially true that a set of A-properties will locally supervene on the set of A-properties itself, regardless of what the A-properties are. They could be the most extrinsic, globally dependent, holistic properties one could imagine, and that local supervenience claim would still hold.

But prevailing usage being what it is, in this appendix I will only speak of "local" as applied to supervenience claims, to avoid confusion. To refer to the "local" or the "individualistic" or the slightly-more-than-intrinsic *properties* on which social properties are often taken to depend, I will speak of "L-properties" rather than "local properties."

A common way of formalizing local supervenience is as "weak local supervenience," as defined by Jaegwon Kim:

> (WLS) *A*-properties *weakly locally supervene* on *B*-properties if and only if for any possible world *w* and any objects *x* and *y* in *w*, if *x* and *y* are *B*-indiscernible in *w*, then they are *A*-indiscernible in *w* (Kim 1984: 163).

Applying this definition to the present case, social properties weakly locally supervene on L-properties if and only if for any possible world w and any entities x and y in w, if x and y are L-indiscernible in w, then they are socially indiscernible in w. Two objects are L-indiscernible or socially indiscernible if and only if they are exactly alike with respect to every L-property or every social property, respectively.

Supervenience has come under criticism in recent years. Many philosophers have begun to doubt its utility for capturing the intuitive strength of dependence claims involving social properties. One reason is that the field has become hair-splittingly technical. There are now so many versions of supervenience (weak, strong, local, global, regional, multiple-domain, etc.) that it is unclear which interpretation of supervenience is the appropriate one to use, if any. Second, a number of philosophers have grown skeptical that supervenience is sufficient to capture the "dependence" of one set of properties on another. It is common to note that supervenience is only a modal relation, involving necessary co-variation of properties, while true dependence claims would seem to need more.

For some uses of supervenience, these are indeed problems. For our purposes, however, they are actually an advantage. If we wanted to *defend* the ontological dependence of social properties on L-properties with a supervenience claim, we would have to show two things: (i) that the supervenience claim was true and (ii) that the demonstrated supervenience claim was sufficient to capture the intuitive dependence claim. To *deny* dependence, however, the case is strongest if we can successfully deny even the weakest of the conditions for dependence to hold. To demonstrate the failure of dependence with a failure of supervenience, it does not matter that supervenience is not sufficient for dependence. What matters is that some form of supervenience is *necessary* for dependence. And while it is possible to deny that the strongest forms of supervenience are necessary for dependence, it can hardly be denied that the weaker forms are.

As I mentioned in the body of the paper, it is widely acknowledged that the local supervenience of social properties on L-properties fails, following Currie's point that Gordon Brown's being Prime Minister does not depend on his own L-properties. This intuitive point is correct. However, it turns out that if we stick to the most commonly used form of local supervenience, it fails on a technicality, rather than for the intuitive reasons people have taken it to. This threatens to mangle the case for local supervenience failure. By shoring up the definition of local supervenience, however, we can highlight the underlying reasons it fails. (For a related discussion of these points, see also Epstein 2009, 210–212.)

To bring out the problem, it is useful to consider a social property of a group, rather than a property of an individual person. Suppose we wish to assess whether *being a freshman class* (let us call that property F) supervenes locally on L-properties, interpreting local supervenience as (WLS) and employing a standard supervenience test. To do so, we take an actual case of a freshman class in context c_1, say the current one at Tufts. Let us call the collection of the members of the Tufts freshman class M. For the doppelganger case, suppose there is a collection of people L-indiscernible from M in some context c_2; call that sum N. In c_2, however, suppose that those individuals' parents have acted differently, failing to enroll them in college at all. The individuals in N are then not members of the Tufts freshman class, or any freshman class at all. Thus F does not hold of N. As stated, the scenario looks as though it demonstrates dependence failure. However, it does not in fact successfully make the same point.

It is important to be careful about which objects we are tracking across possible worlds—i.e., which objects x and y we are tracking to see whether setting their L-properties to be the same guarantees that they have the same social properties. Intuitively, we want to consider whether the social properties of the freshman class are fixed by its individualistic properties. So we want to compare the freshman class in c_1 with the freshman class in c_2, fixing them to have the same individualistic properties. But of course, there is no freshman class in c_2, so this cannot work. Without taking care, the supervenience case does not even get set up. This is the reason I set the case up in terms of mere collections of individuals M and N: across contexts c_1 and c_2, we fix the L-properties of the collections of individuals to be indiscernible, and then evaluate whether they have the same social properties.

But in taking care on this tracking issue, on closer inspection even properties that intuitively should locally supervene will fail to, if we use (WLS). In fact, (WLS) will fail even if we only consider objects in the local context, not even considering the doppelganger context at all. Suppose 10 of us choose to form a literary group. Let us stipulate that membership in the group is a matter only of mutual agreement among us. As with the case above, one of the entities in the local environment is the collection of the ten of us. The problem, however, is that the properties *being a literary group* and *being a freshman class* do not hold of mere collections. This is a point that is familiar in discussions of the metaphysics of "coincident objects." A statue, for instance, may have aesthetic properties that the lump of clay constituting it does not, and a person may have mental properties that the lump of tissue constituting her may not. (See Zimmerman 1995; Fine 2003; Koslicki 2004; Bennett 2004). As applied to social groups, this problem is even more apparent and yields local supervenience failure not only for properties like *being a freshman class,* that intuitively should fail, but it also yields local supervenience failure for any number of properties like *being a book group* that seem as though they should supervene locally on L-properties. What is needed is an understanding of local supervenience that has the ability to preserve the intuition that certain properties of groups *do* supervene locally on L-properties.

One way of solving the problem is to choose the properties judiciously, to show the expected failure and success of weak local supervenience using (WLS). The (WLS) is satisfactory if we ignore property F, *being a freshman class,* applied either to the Tufts freshman class (to which F holds necessarily) or to sum M (to which F does not hold), but rather consider applying to M the related property *coinciding with the membership of a freshman class.* Call this property F'. F' is a social property, one that does apply to M, and that fails to supervene on the L-properties of M. This, then, is an example of a social property (albeit an unusual one) that does not get retained in the indiscernible counterfactual case. Inasmuch as F' is a social property, the argument demonstrates the failure of local supervenience of social properties on L-properties. It seems, however, that we ought to be able to put on a stronger case than this.

A different approach is to follow Zimmerman, Bennett, and others in weakening local supervenience to be less stringent about the entities to which properties are taken to apply. The idea of "coincident-friendly local supervenience", roughly speaking, is that when we assess the social properties of any L-indiscernible pair, we do not only see if *that pair* is socially indiscernible. Rather, we look around the domain for other pairs of objects that coincide with and are L-indiscernible from the respective members of the original pair. In other words, in assessing whether property F applies to M in context c_1 and a sum N in c_2, we do not stop when we see that F does not hold of sums at all, and hence not to M or N. Rather, we look for other entities that coincide with M and N, respectively, and see if they have property F. If there is *any* pair of such entities, then F is regarded as holding for that pair.

On this slightly weaker but still plausible interpretation of local supervenience, supervenience failure does not follow just from the fact that ordinary social objects exist and have properties that do not apply to sums. Instead, it fails for more intuitive reasons, and in fact it shows how properties like F, in addition to properties like F', also fail to supervene on L-properties.

For property F, *being a freshman class,* suppose that in the doppelganger case there are no universities in c_2, so there is no freshman class there at all. (As always, N remains L-indiscernible from the actual collection of individuals M.) In the actual case, there is an entity coinciding with M and L-indiscernible from it that has property F. But in c_2, although there is an entity that is L-indiscernible from M, no entity has property F. Local supervenience thus fails. This conforms to the intuitive point. If *holding the office of President* does not supervene locally on L-properties, then neither does *being a president;* and if *coinciding with the membership of a freshman class* does not supervene locally on L-properties, then neither does *being a freshman class.* Hence the set of social properties that fail to supervene locally on L-properties does not only include the peculiar membership properties of individuals or groups but also the more intuitive properties that apply to social entities as well. This conception of supervenience also avoids forcing supervenience failure for those social properties that should locally supervene on L-properties, such as *being a literary group.*

Despite some technical hiccups, the intuitive point thus remains unchanged: social properties fail to depend ontologically on the "local" properties of the individuals or groups that bear them. Since social properties may depend on the wider population, they can change or fail to hold even if those wider factors do not interact causally with the bearer. Again, we often make the choice not to model parts of the dependence base of any given property. But it is not clear why we should systematically ignore all nonlocal factors on which social properties ontologically depend and only model the local ones.

NOTES

1. See, for instance, Tesfatsion 2003; Epstein 2005; Axtell 2006; Samuelson and Macal 2006.
2. Lotka 1925; Volterra 1926; Anderson and May 1992.
3. Epstein and Axtell 1996.
4. Typical assumptions of representive agent models, as Kirman 1992 points out, are that "all individuals should have identical homothetic utility functions (that is, ones with linear Engel curves); or that all individuals should have homothetic utility functions, not necessarily identical, but that the relative income distribution should be fixed and independent of prices" (120; cf. Lucas 1975; Woodford 2003). Some nascent analytic heterogeneous agent models are discussed in Heer and Maussner 2005. For criticism of the homogeneity of representative agents, see Kirman 1992; Hoover 2006.
5. Mortveit and Reidys 2008 limit cellular automata to synchronously updating systems, though the term 'cellular automaton' is sometimes applied more broadly to asynchronously updating ones.
6. Other less general structures that fall under generic graph-dynamical systems include Boolean networks, graph automata, concurrent transition systems, recurrent Hopfield networks, etc. Different agent-based modeling software packages (e.g., Swarm, Repast, Netlogo) provide for many of these variations.

7. Barrett and Reidys 1999; Barrett et al. 2000; Barrett et al. 2002; Barrett et al. 2006; Mortveit and Reidys 2008.
8. This is a minor variant on the presentation in Barrett et al. 2006.
9. North and Macal 2007, 25. Below I criticize this picture.
10. Minar et al. 1996.
11. This does not entail that the properties of the high-level agent are reducible to those of the constituent agents. In a similar vein, Kevin Hoover has discussed macroeconomic expectation variables that are determined by the properties of individuals and yet are not reducible to the expectations of individuals (Hoover 2009, 406). But, as Hoover stresses, that macroeconomic variable "does not possess causal capacities of the type championed by Cartwright 1989 that can be carried from one context to another. It is not cause or effect, but a summary statistic without causal efficacy." My thanks to an anonymous reviewer for pointing out this connection.
12. Historically, the ontological issues pertaining to the dependence of social properties on the properties of individual people were mixed up with issues pertaining to reductive explanation. Recent work has been careful to separate ontological individualism from "explanatory individualism." It is common for contemporary anti-reductionists to endorse ontological individualism and deny explanatory individualism (cf. Lukes 1968; Pettit 2003).
13. See Kim 1984, 1987; McLaughlin 1995; McLaughlin and Bennett 2005. In Epstein 2009, I discuss supervenience claims in connection with ontological individualism.
14. That social properties globally supervene on individualistic properties is a weak claim. In Epstein 2009, I weaken it further and argue that it nonetheless fails.
15. Elsewhere I have written about these in detail. See also Epstein 2008. Here I will characterize them intuitively rather than technically, and focus on how they are particularly destructive when combined with one another.
16. We have to tread carefully in applying troubles with individualism to agent-based models in the social sciences. As I will point out, the factors that agent-based models incorporate are often not restricted to those factors that can plausibly be understood to be individualistic. So in arguing against agent-based modeling, we cannot just apply the individualistic fallacies directly. However, a version of the same fallacies will nonetheless extend to agent-based modeling.
17. Cf. Teller 2001. I am grateful to an anonymous reviewer for drawing my attention to this reference and example.
18. This is because the objects being modeled are intrinsically individuated, i.e., their identity properties, such as *being a pot* or *being a block of ice,* depend only on intrinsic features of the object itself.
19. This distinction is sometimes overlooked because equations and functions typically do not distinguish the different uses of the equals sign or the function symbol, and hence "constitutive equations" are not visibly distinct from "dynamic" and other equations.
20. The idea of a property being "local" I mean to be somewhat looser than a property being "intrinsic." The reason I distinguish a property being local from a property being intrinsic is that individualistic properties are arguably not intrinsic properties of individual people. But they are, to some extent, local properties (e.g., that I am standing on a stone floor is plausibly an individualistic property of me; cf. Epstein 2009). For intuitive purposes, though, it is fine to use "local" and "intrinsic" interchangeably, until we want to be precise about talking about the relation between social properties and individualistic properties. Intuitively, an intrinsic property of an object is a property that holds or fails to hold regardless of what other objects there are in the world. A local property of an object is one that holds or fails to hold regardless of what the world is like at a spatiotemporal distance from the object.
21. I am grateful to two anonymous reviewers for pressing me on this. One anonymous reviewer usefully called this a "Markov condition for ontological dependence."
22. Oppenheim and Putnam 1958.
23. See Kim 2002 for an attempt to weaken these conditions, and Epstein manuscript (forthcoming) for reasons this weakening fails.
24. Virchow 1860.
25. Fine 2003; Bennett 2004; Koslicki 2004. On the constitution of social groups, see Uzquiano 2004.
26. For present purposes, ontological dependence may be understood as supervenience, so the question as to what the property depends on can be understood as the question of the property's supervenience base. In (Epstein 2009) I present a more detailed discussion of the varieties of supervenience involved and their relation to ontological dependence in the case of social properties.
27. Wimsatt 1994, Epstein manuscript (forthcoming).
28. A way to specify this more precisely: taking facts to be property exemplifications, a pair <M,P>, where M is a set of n-level entities and P is a set of n-level properties, is *simply dependent* if and only if there is a pair <M', P'> such that: (i) M' is a set of (n-1)-level entities and P' is a set of (n-1)-level properties; (ii) The entities in M are exhaustively composed by entities in M'; (iii) P supervenes locally on P'; and (iv) For every entity m in M, m is locally individuated by the P' properties of the M' constituents of m, (i.e., identity conditions for m are local n-1-level properties). This definition involves a levels hierarchy of properties as well as a compositional hierarchy of entities.
29. The base ontology need not be unique.
30. Certainly the choices of target, base, causal, modeled-base, and modeled-causal ontologies are not explicit parts of model-construction, but it is not implausible to see them as implicit.
31. This does not require that the macroentities be reducible to microentities. One way of putting the reason simple dependence does not entail reducibility is that there may not be type-identities between the levels. Another way of putting this point is that simple dependence may hold because a global supervenience relation holds between entities at n and n-1, which may be compatible with nonreducibility. (Although some philosophers have recently argued that supervenience entails reducibility. See, for example, Kim 2005).
32. Typically, this reflects the common Galilean idealization in modeling of treating open systems as if they were closed or isolated and neglecting factors that have little causal influence (cf. Strevens 2004; Weisberg 2007)
33. The idea of a system being "well-behaved" is intuitive, but difficult to analyze. If a system is highly volatile, and develops along many radically different paths in different scenarios, interacting with radically different parts of the world in each, then it is not well behaved.
34. It is easiest to describe changes to be made to the modeled ontology as being chosen by the modeler or some external observer. But in fact this is not a necessity: an "external perspective" can be taken computationally, to modify the dynamical system being modeled. For instance, a governing process might be triggered when some entity that was expected to stay in the center of a cellular automaton runs off to the edge, and restart the simulation with a bigger grid. Or consider a model with a "synthetic" population, having agents generated by some statistical algorithm. A process might automatically detect

that people out in the suburbs, who were expected to be superfluous, actually will materially affect the end results and then generate an expanded population and restart the run.

35. Along with the nonlocal dependence of macro-properties, the causal interactions may also be "wide." cf. Yablo 1997.
36. If macro-properties do need to be included in models, then presumably micro-foundational responses to the Lucas critique (Lucas 1976) need to be revisited.
37. On this issue, see Epstein and Forber, manuscript (forthcoming). Some critics of representative agent models have pointed out that the properties of representative agents are simply macroscopically measured parameters divided by the population, which they have argued are macroscopic properties in microscopic clothing. Using macroscopic properties to set the parameters of agents may be similar. On the other hand, some philosophers have argued that the traditional injunctions against using desired results in hypothesis generation (i.e., against the use of "old data") are overstated. Which may make the use of macrodata in the setting of parameters of agents acceptable and also may allow us to count these parameters as genuinely micro-properties, even though they are determined on the basis of macro-properties.
38. This applies even to recent approaches to multiscale modeling, such as the methods described in E et al. 2007

REFERENCES

Anderson, Roy M., and Robert M. May (1992). *Infectious Diseases of Humans: Dynamics and Control.* Oxford: Oxford University Press.

Axtell, Robert (2006). *Agent-Based Computing in Economics,* http://www2.econ.iastate.edu/classes/econ308/tesfatsion/ACEIntro.RAxtell.pdf, accessed 15 September 2008.

Barrett, Chris, and Christian Reidys (1999). "Elements of a Theory of Simulation I: Sequential CA over Random Graphs." *Applied Mathematics and Computation* 98: 241–259.

Barrett, Chris, Henning Mortveit, and Christian Reidys (2000). Elements of a Theory of Simulation II: Sequential Dynamical Systems. *Applied Mathematics and Computation* 107(2–3): 121–136.

Barrett, Chris, Stephen Eubank, Madhav Marathe, Henning Mortveit, and Christian Reidys (2002). "Science and Engineering of Large Scale Socio-Technical Simulations." Proceedings of First International Conference on Grand Challenges in Simulations at San Antonio, Texas.

Barrett, Chris, Keith Bissett, Stephen Eubank, Madhav Marathe, and V.S. Anil Kumar (2006). "Modeling and Simulation of Large Biological, Information, and Socio-Technical Systems." Proceedings of the Short Course on Modeling and Simulations of Biological Networks, at AMS Lecture Notes.

Bennett, Karen (2004). "Spatio-Temporal Coincidence and the Grounding Problem." *Philosophical Studies* 118: 339–371.

Cartwright, Nancy (1989). *Nature's Capacities and Their Measurement.* Oxford: Clarendon Press.

Currie, Gregory (1984). "Individualism and Global Supervenience." *British Journal for the Philosophy of Science* 35: 345–358.

E, Weinan, Bjorn Engquist, Xiantao Li, Weiqing Ren, and Eric Vanden-Eijnden (2007). "Heterogeneous Multiscale Methods: A Review." *Communications in Computational Physics* 2(3): 367–450.

Epstein, Brian (2008). "When Local Models Fail." *Philosophy of the Social Sciences* 38: 3–24.

Epstein, Brian (2009). "Ontological Individualism Reconsidered." *Synthese* 166(1): 187–213.

Epstein, Brian (Forthcoming). "Science without Levels." Unpublished manuscript.

Epstein, Brian, and Patrick Forber (Forthcoming). "The Perils of Tweaking." Unpublished manuscript.

Epstein, Joshua M. (2005). "Remarks on the Foundations of Agent-Based Generative Social Science." Washington, DC: The Brookings Institution, Brookings CSED Working Paper #41.

Epstein, Joshua M., and Robert Axtell (1996). *Growing Artificial Societies: Social Science from the Bottom Up.* Cambridge: MIT Press.

Fine, Kit (2003). "The Nonidentity of a Material Thing and its Matter." *Mind* 112: 195–234.

Heer, Burkhard, and Alfred Maussner (2005). *Dynamic General Equilibrium Modelling.* New York: Springer.

Hoover, Kevin (2006). "A NeoWicksellian in a New Classical World: The Methodology of Michael Woodford's Interest and Prices." *Journal of the History of Economic Thought* 28 (2): 143–149.

Hoover, Kevin (2009). Microfoundations and the Ontology of Macroeconomics. In *Oxford Handbook of Philosophy of Economics*, eds. Harold Kincaid and Don Ross. Oxford: Oxford University Press, 386–409.

Kim, Jaegwon (1984). "Concepts of Supervenience." *Philosophy and Phenomenological Research* 45: 153–176.

Kim, Jaegwon (1987). "'Strong' and 'Global' Supervenience Revisited." *Philosophy and Phenomenological Research* 45: 153–176.

Kim, Jaegwon (2002). "The Layered Model: Metaphysical Considerations." *Philosophical Explorations* 5(1): 2–20.

Kim, Jaegwon (2005) *Physicalism, or Something Near Enough.* Princeton: Princeton University . Press.

Kirman, A. (1992). "Whom or What Does the Representative Agent Represent?" *Journal of Economic Perspectives* 6: 117–136.

Koslicki, Kathrin (2004). "Constitution and Similarity." *Philosophical Studies* 117: 327–364.

Lotka, Alfred J. (1925). *Elements of Physical Biology.* Baltimore: Williams and Wilkins.

Lucas, Robert E. (1975) "An Equilibrium Model of the Business Cycle." *Journal of Political Economy* 83: 1113–1144.

Lucas, Robert E. (1976). "Econometric Policy Evaluation: A Critique." In *The Phillips Curve and Labor Markets*, eds. Karl Brunner and Allan H. Meltzer. Amsterdam: North-Holland, 19–46.

Lukes, Steven (1968). "Methodological Individualism Reconsidered." *British Journal of Sociology* 19: 119–29.

McLaughlin, Brian P. (1995). "Varieties of Supervenience." In *Supervenience: New Essays*, eds. Elias Savellos and Umit Yalcin. Cambridge: Cambridge University Press, 16–59.

McLaughlin, Brian P., and Karen Bennett. (2005)."Supervenience." *Stanford Encyclopedia of Philosophy,* http://plato.stanford.edu/entries/supervenience, accessed 15 September 2008.

Minar, Nelson, Roger Burkhart, Chris Langton, and Manor Askenazi (1996). The Swarm Simulation System: A Toolkit for Building Multi-Agent Simulations, http://swarm.org/images/b/bb/MinarEtAl96.pdf, Accessed 15 September 2008.

Mortveit, Hennig, and Christian Reidys (2008). *Introduction to Sequential Dynamical Systems.* New York: Riedel.

North, Michael, and Charles Macal (2007). *Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation.* New York: Oxford University Press.
Oppenheim, Paul, and Hilary Putnam (1958). "Unity of Science as a Working Hypothesis." In *Minnesota Studies in the Philosophy of Science.* Minneapolis: University of Minnesota Press, 3–36.
Pettit, Philip (2003). "Groups with Minds of Their Own." In *Socializing Metaphysics*, ed. Frederick Schmitt. Lanham, MD: Rowman and Littlefield, 167–193.
Ruben, David-Hillel (1985). *The Metaphysics of the Social World.* London: Routledge and Kegan Paul.
Samuelson, Douglas, and Charles Macal (2006). "Agent-Based Simulation Comes of Age." *Operations Research Management Science (OR/MS) Today* 33(4):34–38.
Schelling, Thomas C. (1978). *Micromotives and Macrobehavior.* New York: W. W. Norton.
Strevens, Michael (2004). "The Causal and Unification Account of Explanation Unified –Causally." *Noûs* 38: 154–76.
Teller, Paul (2001). "Twilight of the Perfect Model Model." *Erkenntnis* 55(3): 393–415.
Tesfatsion, Leigh (2003). "Agent-Based Computational Economics." Iowa State University Economics Working Paper, Ames IA.
Uzquiano, Gabriel (2004). "The Supreme Court and the Supreme Court Justices: A Metaphysical Puzzle." *Noûs* 38(1): 135–153.
Virchow, Rudolf Ludwig Karl (1860). *Cellular Pathology: As Based Upon Physiological and Pathological Histology.* Trans. F. Chance. London: John Churchill.
Volterra, Vito (1926). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Memoria della Regia Accademia Nazionale dei Lincei* 6(2):31–113.
Weisberg, Michael (2007). "Three Kinds of Idealization." *Journal of Philosophy* 104(12): 639–659.
Wimsatt, William (1994). "The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets." *Canadian Journal of Philosophy* Supplement 20: 207–274.
Woodford, Michael (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton: Princeton University Press.
Yablo, Stephen (1997). "Wide Causation." *Philosophical Perspectives* 31(11): 251–281.
Zimmerman, Dean (1995). Theories of Masses and Problems of Constitution. *Philosophical Review* 104: 53–110.

7 Scientific Models, Simulation, and the Experimenter's Regress

Axel Gelfert

INTRODUCTION

In this paper, I analyze the question of whether computer simulation is, in any special way, affected by what has variously been called the "experimenter's regress" (Collins 1985) or "data-technique circles" (Culp 1995). Such a regress, it has been argued, may obtain when the only criterion scientists have for determining whether an experimental technique (or simulation) is 'working' is the production of 'correct' (i.e., expected) data. It may seem plausible to assume that techniques of computer simulation are especially prone to such regress-like situations, given that they are further removed from nature (in ways to be specified) than traditional experimentation. In public perception, too, there appears to be a gap between the trust that is placed in the experimental success of science, as opposed to its use of computer simulation methods (e.g., in predicting global climate change).

The rest of this paper is organized into eight sections. The first section summarizes the main idea of the *experimenter's regress*, as developed by Harry Collins (1985) on the basis of a case study in experimental astrophysics. The following section then addresses the question of whether computer simulation can properly be thought of as a form of 'experimenting on theories'. The section "Formulating the 'Simulationist's Regress'" identifies three clusters of scientific questions, where one might expect regress-like situations to arise in connection with computer simulation; this is followed by a distinction, in "Anatomy of a Regress", between what I call a 'software' and a 'hardware' aspect of replicability, both of which might be thought of as contributing to worries about a potential 'simulationist's regress.' I then consider the 'Response from Robustness' according to which the experimenter's regress can be dissolved whenever independent measurement techniques generate sufficiently 'robust' data. However, as I argue in "Models as Targets", in the case of simulation, this standard response is not typically available, since in actual scientific practice, computer simulation studies do not always have real systems as its targets; what is being 'simulated' are (sometimes quite abstract) mathematical models. The penultimate section, "Robustness, Invariance, and Rigorous Results" develops a richer notion