14

MODELS ARE STUPID, AND WE NEED MORE OF THEM

Paul E. Smaldino

> All social science research must do some violence to reality in order to reveal simple truths.
>
> *(Lazer & Friedman, 2007)*

Despite numerous efforts extolling the virtues of formal modeling (Epstein, 2008; Schank, 2001; Smith & Conrey, 2007; Marewski & Olsson, 2009; Farrell & Lewandowsky, 2010; Weinhardt & Vancouver, 2012; Smaldino, Calanchini, & Pickett, 2015), there remains widespread resistance among social and behavioral scientists to adopt formal modeling in their general research approach. In addition to the technical challenge posed by the mathematical and programming skills required to understand and develop models, a common point of resistance appears to stem from the perception of models as crude, overly simplistic, and unrealistic. The conclusion is that models are largely useless as anything but a formal exercise, and unnecessary for most scientists to engage with.

Rather than argue against this perception, I enthusiastically embrace the perspective of the resistance, at least in part. Models are, by and large, stupid. My point of contention is with the conclusion that stupid models are not useful. Quite the contrary. Stupid models are extremely useful. They are useful because humans are boundedly rational and because language is imprecise. It is often only by formalizing a complex system that we can make progress in understanding it. Formal models should be a necessary component of the behavioral scientist's toolkit. Models are stupid, and we need more of them.

We Are Stupid

Down to the very name of our species, *Homo sapiens*, we humans love to emphasize our intelligence relative to other species. We can certainly solve many complicated problems. And yet we are often very stupid animals who make foolish choices. This isn't a raw failing on our part. We are limited beings, with finite resources with which to compute a coarse model of our world and with which to invent options and evaluate their consequences. Moreover, our world, and the ecological and social environments in which we find ourselves, are changing rapidly, far too rapidly for our brains to possibly adapt via genetic evolution. We do the best we can.

Humans appear to have particular difficulty understanding complex systems. Mitch Resnick, in his book *Turtles, Termites, and Traffic Jams*, details his experiences teaching gifted high school students about the dynamics of complex systems using artificial life models (Resnick, 1994). He showed them how organized behavior could emerge when individuals responded only to local stimuli using simple rules, without the need for a central coordinating authority. Resnick reports that even after weeks spent demonstrating the principles of emergence, using computer simulations that the students programmed themselves, many students still refused to believe that what they were seeing could *really* work without central leadership.

We who study complex systems for a living may feel a certain smugness here. The average person may have difficulty understanding the forces that drive behavior, we think, but through our powerful intellects, our education, and our hefty experience pondering the deep mysteries, *we* can trust our intuition when it comes to understanding the psychological and social forces that make people do what they do. Unfortunately, my own experience working with complex systems and working among complexity scientists suggests that we are hardly immune to such stupidity. Indeed, even seemingly simple puzzles can pose a challenge.

Consider the case of Marilyn Vos Savant and the Monty Hall problem. Vos Savant, famous for her record high score on standard IQ tests, has written a weekly puzzle column in *Parade Magazine* since 1986. In 1990, she wrote about a puzzle commonly known as the Monty Hall problem. The problem goes as follows. You are on a game show and given the choice to open one of three doors. Behind one of the doors is a fabulous cash prize, and behind the others, goats (the assumption is that no one would prefer the goats to the cash). You choose a door, say Door #1. The host, who knows where the cash really is, opens one of the other two doors, say #3, and shows you a goat behind it. The host now offers you the option to switch to Door #2. The question is whether it is to your advantage to do so.

The answer is that, although you are never guaranteed to be correct, you should probably switch. The cash is twice as likely to be behind Door #2 instead of Door #1. This is not an easy result for most people to wrap their heads around, though it follows quite definitively from the assumptions of probability theory (if you are in doubt of the problem's trickiness, I suggest that you pose it the next time you are at a dinner party). Strikingly, Vos Savant's answer was challenged not only by lay readers, but also by many with advanced mathematical training. Indeed, she received many letters from professional mathematicians insisting that she was mistaken, even after she published a follow-up column with a detailed proof. The letters were often written in a smug, knowing tone; Vos Savant details many of these in an article posted to her website (http://marilynvossavant.com/game-show-problem/). One, written after the publication of the follow-up column and signed by a Georgetown University professor, reads:

> You are utterly incorrect about the game show question, and I hope this controversy will call some public attention to the serious national crisis in mathematical education. If you can admit your error, you will have contributed constructively towards the solution of a deplorable situation. How many irate mathematicians are needed to get you to change your mind?

It is my belief that the widespread inability to grasp the solution to the Monty Hall problem stems from a failure to properly model the scenario. You should switch doors because regardless of which door you picked initially, the host can *always* show you one with a goat. Being shown a goat therefore has no bearing on the probability that your initial choice was correct. Since that probability is 1/3, there is a 2/3 chance that you were wrong and the cash is behind the remaining door. Thus, two out of three times, switching is the right move. The common intuition that the choice is instead a 50/50 split between two options is erroneous.

Readers of this chapter are likely to be interested in social behaviors and their underlying psychological mechanisms. These systems tend to be quite a bit more complicated than a simple game show problem. This should concern us. Being an expert does not inoculate us from the failure of our limited imaginations, which evolved to solve problems quite different from those of interest to behavioral scientists. We could use some help.

Models to the Rescue?

I am, of course, going to argue that we should turn to models, and particularly *formal* models, for help. Specification of a formal model delineates the parts of a system and the relationships between those parts, allows us to examine the logical conclusions of our assumptions, and as a byproduct, examine the appropriateness of those assumptions in the first place. But first, I need to take a brief detour, because when it comes to explaining any behavior, the first question we need to ask is: *What are we talking about*?

Articulating a System and Its Parts

As behavioral and social scientists, we want to understand some system related to individual or social behavior. Maybe we are interested in how social identity manifests when individuals feel threatened, or how individuals coordinate in joint activities, or how racially charged language is interpreted by individuals of different racial and socioeconomic backgrounds. These examples obviously represent a miniscule selection from among the questions we might ask. The important thing to note is how each question is subject to myriad interpretations. What aspect of the behavior are we interested in, specifically? Are we interested in the neurophysiology of joint attention, down to the way neural spike trains inform action programs? Or are we more interested in a "higher" level of organization, perhaps one in which we can ignore physiology and instead simply consider the temporal relationships between individually designated behavioral units? These are not trivial questions. For any given behavior, there are many questions we can ask related to its development, mechanism, and adaptive function, none of which are obviously favored from a scientific perspective (Tinbergen, 1963).

Once we specify the level of organization and the kind of explanation we are looking for, we still need to do additional work to specify the exact question under investigation. Human beings are complex beings. It's not just that we exist at many levels of organization. Of course, we are made of organs, which are made of tissues, which are made of cells, which communicate using molecules and ions; above the level of the individual, we are enmeshed in local social networks, communities both corporate and categorical, economies, and nations. A further problem arises when we consider that these levels interact—the causal arrows flow both ways (Campbell, 1974; Wimsatt, 1974). The problem is not insurmountable, but needs to be acknowledged. Any explanation of individual and social behavior must necessarily ignore important causal relationships both within and between levels of organization. We must become comfortable with ignoring those relationships, and this comfort is achieved partly through acknowledging their existence.

Part of specifying our research question involves the articulation of the parts of the system of inquiry. Kauffman's 1971 essay still provides the best discussion of this important but overlooked issue. Notice that I do not say that we should specify our question and *then* articulate the parts of the system. The two are parts of a single process. What is our question? To understand joint attention in coordinated behavior, perhaps. But what is our question *right now*? We must decompose the system into explicit parts. We must postulate properties of those parts and the relationships between them. In some sense, this is the essence of all scientific inquiry into behavior. All well-formed scientific research questions concern the properties of parts, the relationships between them, and the consequences of those relationships. The articulation of parts and relationships will necessarily be overly simplified and ignore details of physical reality. But much like a map is only useful because it ignores irrelevant detail, so is a well-formed scientific question useful when it captures only those features of reality most relevant to a useful answer.

To make myself perfectly clear: To ask a scientific question about individual or social behavior, we must specify the parts of a system and the relationships between them. The question at hand may be about the *nature* of these parts or their relationships, and so we may designate a *distribution* of parts or relationships from which to sample, but it amounts to the same thing. The precise specification of parts and relationships is what defines a scientific question and separates it from wishy-washy pseudotheory that is unfalsifiable and distracting (Popper, 1963; Gigerenzer, 1998; Smaldino, 2016).

Building Models, Formal and Otherwise

Let us assume that we have articulated, in words, the parts of our system and the relationships between them. Perhaps we say, as do the adherents of optimal distinctiveness theory (Brewer, 1991; Leonardelli, Pickett, & Brewer, 2010), that individuals have social identities that correspond to different contexts and different levels of inclusivity, and that they express these identities in order to balance internal drives for assimilation and differentiation. The parts are obviously the individuals, each of whom has the property of possessing an array of identities and the ability to express one of these at any given time. The relationships between the parts manifest as perceptions of others' identities, which dictate how individuals update their own expression. The theory suggests how this updating might occur: individuals should express more exclusive identities when their currently expressed identity is very inclusive, and vice versa.

I have just described what is often called a "verbal model." As Epstein (2008, para. 1.2) phrases it, "Anyone who ventures a projection, or imagines how a social dynamic . . . would unfold is running *some* model." Most behavioral and social scientists are quite comfortable with this sort of model. However, look closer. You'll see that the parts of the system are not particularly well articulated, and neither are their relationships. What does it mean to possess an identity, let alone an array of them? How do individuals choose between their identities when it comes time to express them? Is the expression of a new identity costly, perhaps in terms of time or social capital? How do individuals take stock of the identities of their fellows? Are their perceptions accurate? Are all identities equally easy to perceive? There are additional related questions as well, concerning the nature of system. Where do identities come from, and how might an individual gain a new identity or lose an existing one? What is the adaptive function of expressing an identity in the first place, since, to be preserved, identities must serve some purpose other than internal contentment?

This is not to pick on optimal distinctiveness theorists. Social psychology, and the social and behavioral sciences more generally, are replete with similar cases. This is the limitation of verbal models. They are often a good way to begin an inquiry when the available evidence suggests only some broad type of relationship that might be further refined. The danger with most verbal models is that there are many ways to specify the parts and relationships of a system that are consistent with such a model. Scientific inquiry stalls when data is used to simply support rather than refine a verbal model. Because many different data sets are consistent with a vague verbal model, researchers using such techniques risk lapsing into positing theories that are, by and large, unfalsifiable (Popper, 1963; Gigerenzer, 1998).

The articulation of the parts of a system and the relationships between them always involves incurring some violence upon reality. Science is an iterative process, and pragmatically, we must ignore some details about complexity and organization to make any headway. That said, it's not a terrible goal to try and be a bit more precise. This is where formal modeling comes in. A formal model instantiates the verbal model as a collection of mathematical relationships and/or algorithmic processes. Rather than saying an individual has something *like* an array of social identities, we can model an individual as a computation object that has *precisely* an array of social identities, which in turn might be modeled as simple numerical values for the sake of comparisons between individuals. My colleagues and I have made models of this type (Smaldino, Pickett, Sherman, & Schank, 2012; Smaldino & Epstein, 2015). More than anything, we have learned that we have a long way to go in understanding the nature and social significance of social identity.

To paraphrase Gunawardena (2014), a model is a logical engine for turning assumptions into conclusions. By making our assumptions explicit, we can clearly assess their implied conclusions. These conclusions will inevitably be flawed, because the assumptions are ultimately incorrect, or at least incomplete. By examining *how* they differ from reality, we can refine our models, and thereby refine our theories, and so gradually we might become *less* wrong (Wimsatt, 1987; Schank, May, & Joshi, 2014; Smaldino et al., 2015). Making formal models of the systems we study is the *only* way to make this possible.

A Brief Note on Statistical Models

When I talk about formal models, I am primarily talking about models whose purpose is to elucidate the mechanisms underlying psychological and behavioral phenomena. Another category of formal model, more familiar to many readers, I'm sure, is the type of model often used in statistical analysis, such as a path model or a linear model. Statistical models are both important and limited, and therefore worth commenting upon, but as they are not my focus here, I will keep my discussion of them brief.

Statistical analyses are necessary and often well-motivated, but we should never forget that they too have models at their core. The generalized linear model, the work horse of the social sciences, models data as being randomly drawn from a distribution whose mean varies according to some parameter. The linear model is so obviously wrong yet so useful that the mathematical anthropologist Richard McElreath has dubbed it "the geocentric model of applied statistics," in reference to the Ptolemaic model of the solar system that erroneously placed the Earth rather than the Sun at the center, but nevertheless produced accurate predictions of planetary motion as they appeared in the night sky (McElreath, 2015). Such models usually assume that one's data are generated by randomly sampling from some distribution—perhaps a Gaussian distribution whose mean tracks some conditional variable. These models are terrifically important in establishing relationships between variables in empirical data sets, and thus for guiding the development of increasingly strong theories. However, many of these models say little about the processes that *actually* generated the data, or about the *mechanistic* nature of relationships between variables. This is the domain of the kinds of formal models I am principally discussing in this chapter. Such models, if sufficiently precise, may utilize data for validation and calibration (e.g., Schank, 2008; Moussaïd, Helbing, & Theraulaz, 2011; Hills, Jones, & Todd, 2012), but this is not strictly necessary for such models to be useful (Wimsatt, 1987; Bedau, 1999; Epstein, 2008; Gunawardena, 2014).

Models Are Stupid

A common objection to formal modeling in the behavioral and social sciences is that they are grossly unrealistic. This is, in general, quite correct. Formal models are often fantastically unrealistic. They ignore huge swaths of reality, including details of individual behavior and environmental complexity. However, framing this fact as a downside is a serious error, particularly if the alternative is to rely instead on verbal models. Verbal models can appear superior to formal models only by employing strategic ambiguity (*sensu* Eisenberg, 1984), giving the illusion of understanding at the cost of actual understanding. That is, by being vague, verbal models simultaneously afford many interpretations from among which any reader can implicitly, perhaps even unconsciously, choose his or her favorite. I will illustrate this point with a simple parable.

The Parable of the Cubist Chicken

One evening long ago, when I was an undergraduate student, a friend and I found ourselves waiting in the basement of a theater for a third friend, an actor about to finish his play rehearsal. There was a large collection of LEGOs in the room, and being of a jaunty disposition and not entirely sober, we amused ourselves by playing with the blocks. One of us—precisely who has been lost to memory—constructed an assembly of red, white, black, and yellow blocks and declared, "Look! It's a Cubist chicken!" The other one of us laughed and heartily agreed that it most definitely looked like a Cubist chicken. We were extremely satisfied with ourselves, not only because it was very silly, but because if in fact we *both* understood the design to be a Cubist chicken, then it surely *was* one. We had identified something *true* about our little masterpiece, and had therefore, inadvertently perhaps, created *art*. This is how liberal arts students amuse themselves.

Our conversation moved on to other topics, but we continued to occasionally comment on the Cubist chicken. After some time had passed, our actor friend entered the room. "Check it out," we said, "a Cubist chicken!" Our friend smiled bemusedly and asked us to explain exactly how the seemingly random constellation of LEGOs represented a chicken. "Well," I said, pointing to various parts of the assemblage, "Here is the head. And here is the body and the legs, and here is the tail." "No!" cried my co-conspirator. "That's all wrong. The whole thing is just the head. Here are the eyes, and the beak, and here is the crest," for my friend had envisioned our chicken as a rooster. And thus the illusion of our shared reality was shattered. We thought we had been talking about the same thing. But when more precision was demanded, we discovered we had not.

Stupidity Is a Feature, Not a Bug

As many a late-night dorm room conversation can attest, humans are capable of very elaborate theories about the nature of reality. The problem is that, as scientists, we need to clearly communicate those theories so that we can use them to make testable predictions. In the social and behavioral sciences, the search for clarity can present a problem for verbal models, and can lead to a depressing recursive avalanche of definitions. What is a preference? A preference is a tendency for certain behaviors. What are those behaviors? It depends on the context. What is a context? This can go on for a while.

Formal models provide a means of escape from the recursive abyss. By restricting our discussion to the model system, we can clearly articulate all the parts of that system and the relationships between those parts, leaving nothing out. This generally leaves us with something that, on the surface, often appears to be pretty stupid. What I mean is that not only are all models wrong, as George Box famously noted; they are *obviously* wrong. However, the stupidity of a model is often its strength. By focusing on some key aspects of a real-world system (i.e., those aspects instantiated in the model), we can investigate how such a system would work if, in principle, we really *could* ignore everything we are ignoring. This only sounds absurd until one recognizes that, in our theorizing about the nature of reality—both as scientists and as quotidian humans hopelessly entangled in myriad webs of connection and conflict—*we ignore things all the time*. We can't function without ignoring most of the facts of the world. Our selective attention ignores most of the sensory input that nevertheless innervates our neurons (as indicated by the well-known "cocktail party effect"). This ignorance is fundamentally adaptive; the bounds to our rationality are severe, and dedication of cognitive resources entails balancing benefits and costs. Causal explanations work in much the same way. By ignoring all but the most relevant information, we are able to impose some modicum of order upon the world. Problems arise when we try to communicate our systems for ordering the world, as each of us has decomposed the world into a somewhat different set of parts and relationships.

Formal models solve this problem by systematizing our stupidity, and ensuring that we are all talking about the same thing.

In the following section, I will provide several concrete examples of how seemingly stupid models help scientists do their science. Before doing that, however, it is worth taking a moment to discuss some general ways in which models that are obviously wrong can nevertheless inform our thought. For example, studying computational models of complex systems can help us to build mental models of some emergent phenomena whose dynamics are otherwise difficult to visualize (Nowak, Rychwalska, & Borkowski, 2013), and the process of model construction can illuminate core uncertainties in one's knowledge of a system (Epstein, 2008). The clearest delineation I have found is William Wimsatt's (1987) list of 12 "functions served by false models," with the understanding that *all* models are false. I therefore reproduce this list, with only light editing, in Table 14.1.

TABLE 14.1 Twelve functions served by false models. Adapted with permission from Wimsatt (1987).

(1)	An oversimplified model may act as a starting point in a series of models of increasing complexity and realism.
(2)	A known incorrect but otherwise suggestive model may undercut the too ready acceptance of a preferred hypothesis by suggesting new alternative lines for the explanation of the phenomena.
(3)	An incorrect model may suggest new predictive tests or new refinements of an established model, or highlight specific features of it as particularly important.
(4)	An incomplete model may be used as a template, which captures larger or otherwise more obvious effects that can then be "factored out" to detect phenomena that would otherwise be masked or be too small to be seen.
(5)	A model that is incomplete may be used as a template for estimating the magnitude of parameters that are not included in the model.
(6)	An oversimplified model may provide a simpler arena for answering questions about properties of more complex models, which also appear in this simpler case, and answers derived here can sometimes be extended to cover the more complex models.
(7)	An incorrect simpler model can be used as a reference standard to evaluate causal claims about the effects of variables left out of it but included in more complete models, or in different competing models to determine how these models fare if these variables are left out.
(8)	Two false models may be used to define the extremes of a continuum of cases in which the real case is presumed to lie, but for which the more realistic intermediate models are too complex to analyze or the information available is too incomplete to guide their construction or to determine a choice between them. In defining these extremes, the "limiting" models specify a property of which the real case is supposed to have an intermediate value.

(continued)

TABLE 14.1 *(continued)*

(9)	A false model may suggest the form of a phenomenological relationship between the variables (a specific mathematical functional relationship that gives a "best fit" to the data, but is not derived from an underlying mechanical model). This "phenomenological law" gives a way of describing the data, and (through interpolation or extrapolation) making new predictions, but also, because its form is conditioned by an underlying model, may suggest a related mechanical model capable of explaining it.
(10)	A family of models of the same phenomenon, each of which makes various false assumptions, has several distinctive uses: (a) One may look for results which are true in all of the models, and therefore presumably independent of different specific assumptions which vary across models. These invariant results are thus more likely trustworthy or "true." (b) One may similarly determine assumptions that are irrelevant to a given conclusion. (c) Where a result is true in some models and false in others, one may determine which assumptions or conditions a given result depends upon.
(11)	A model that is incorrect by being incomplete may serve as a limiting case to test the adequacy of new, more complex models.
(12)	Where optimization or adaptive design arguments are involved, an evaluation of systems or behaviors which are not found in nature, but which are conceivable alternatives to existing systems, can provide explanations for the features of those systems that are found.

Some (Not So) Stupid Models

Compiling a list of all the interesting and useful models in the sciences is a fool's errand. Let it suffice to say that such a list would be vast. Instead, I want to merely illustrate via a few pointed examples how simple, stupid models can be not only useful, but fundamental to good science. I will start with four well-known examples of models that changed our understanding of basic concepts in the physical, biological, and social sciences. I will then give two examples of how I have used formal models in my own work, focusing on topics that should be of interest to social psychologists: (1) social identity and distinctiveness and (2) hypothesis testing and replication.

Newton's Model of Universal Gravitation

In 17th-century Europe, the field of astronomy faced a great challenge. Following the pioneering work of Copernicus and building on the meticulously collected data of Tycho Brahe, Johannes Kepler had definitively showed that not only do the Earth and the other planets revolve around the Sun, their orbital paths describe ellipses rather than perfect circles. It was a great mystery why this should be. Enter Isaac Newton. Newton was not the first person to propose that the heavenly bodies might be attracted to one another with a force that varied with the inverse square of the distance between them, but he was the first to build a model based on that proposition (Gleick, 2004). His model was startlingly simple, consisting of only two objects—the Sun and the Earth (Figure 14.1). The model ignored the Moon as well as the five other known solar planets, not to mention all the celestial bodies that were unknown in Newton's time. The size and topology of the Sun and Earth were also ignored; they were modeled as points identified only by their mass, position, and velocity. Nevertheless, the model's strength lies in its simplicity. By restricting the analysis to only two bodies, the resulting planetary orbit was mathematically tractable. Using a simple rule stating that the force of gravitation was proportional to the product of the objects' masses and inversely proportional to the square of the distance between them, Newton was able to show that the resulting orbits would always take the form of conic sections, including the elliptical orbits observed by Kepler. And because he could show that the same law explained the motion of falling objects on Earth, Newton provided the first scientific unification of the Terrestrial with the Celestial. Newton's theory of Universal Gravitation rested on a model that, to naïve eyes, can easily appear quite stupid. Ultimately, the theory has been shown to be incorrect, and has been epistemically replaced by the theory of General Relativity. Nevertheless, the theory is able to make exceptionally good approximations of gravitational forces—so good that NASA's Moon missions have relied upon them.

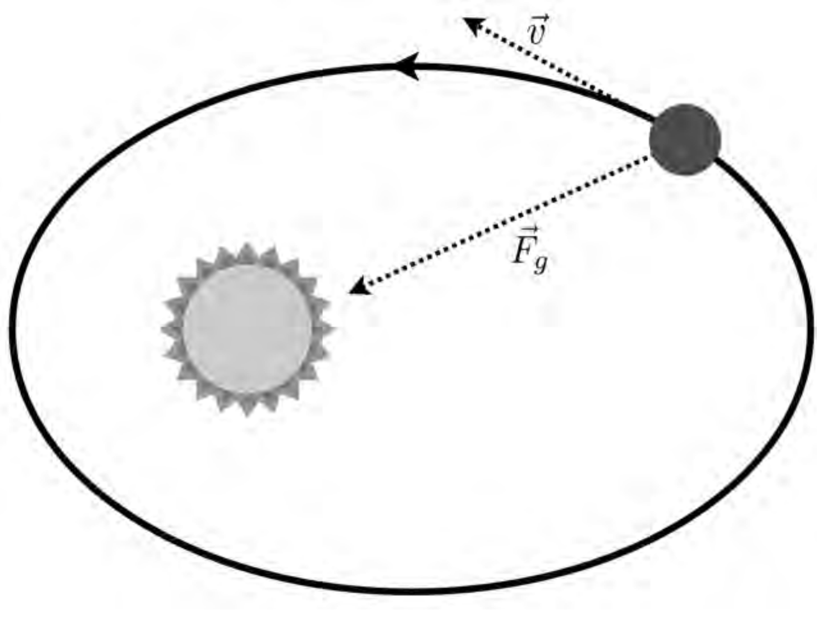

FIGURE 14.1 A graphical representation of Newton's model of planetary gravitation. The Earth has a forward velocity v, which is continuously altered by the gravitational attraction of the Sun, F_g, resulting in an elliptical orbit. In reality, the model is even simpler than implied here, because the Sun and Earth were represented as point masses rather than spheres.

The Lotka-Volterra Model of Predator–Prey Relations

For many years, fur trapping organizations like the Hudson's Bay Company in Canada kept meticulous records on the pelt-producing animals in the regions where they trapped. These records illustrated that linked predator and prey species, like the Canada lynx and the snowshoe hare, tended to have cyclical population levels whose dynamics were tightly correlated. How to explain this? In the early 20th century, Alfred Lotka and Vito Volterra, working independently, applied ideas from the chemistry of autocatalytic reactions to generate a simple model of two interrelated populations, which can be instantiated as a pair of coupled differential equations. This model specifies two animal species: a prey species with a positive rate of growth in the absence of predators, and a predator species with a negative growth rate in the absence of prey. The number of predators negatively influences the number of prey, and the number of prey animals positively influences the number of predators. The model can produce correlated oscillations in the two populations that bear a striking resemblance to data on many predator–prey systems. The model also identifies conditions under which the two growth rates can give rise instead to more stable equilibria as well as yielding complete population collapse—predictions that have since been borne out empirically. However, the model is extremely simplistic. It assumes perfect mixing, so that the probability of a prey animal encountering a predator is simply the relative frequency of predators in the population. It ignores seasonality, circadian cycles, migration, density dependence in the growth rate of the prey species, development, and interactions with other species. Thus, when these features matter, the model may fail to align with empirical fact (Luckinbill, 1973; Berryman, 1992). Nevertheless, the core assumptions of the model often hold. This provides opportunities for extensions and refinements of the model when additional features cannot be ignored. By providing a foundational structure, the Lotka-Volterra model remains one of the core tools for understanding the relationship between predator and prey populations.

Hopfield's Model of Content-Addressable Memory

Memory—the ability to store information for later recall—is a wondrous property of neural networks that makes possible all but the most rudimentary forms of cognition. By the early 1980s, long-term potentiation—the process by which Donald Hebb's theory that "neurons that fire together wire together" occurs—was relatively well described. It was believed that the formation of such associations was intrinsic to more complex forms of memory, such as that by which a person's face is encoded and then later recognized, but the mechanism was unclear. How could a brain possibly use partial information, like an occluded face, to reconstruct information encoded in memory? To begin to answer this question, the biophysicist John Hopfield (1982) constructed a simple model of two-state neurons in a fully connected network. Edge weights were determined by a process of Hebbian learning assumed to have already occurred, so that a number of configurations (or states) of "on" and "off" neurons were encoded in the network, with edges assumed to be bidirectionally symmetrical (i.e., undirected). Using mathematics derived from statistical physics, Hopfield showed firstly that in such a system, encoded states would be stable, and secondly that if initialized in a non-encoded state, the network would self-organize into the encoded state that most closely matched the initialized state. In other words, he had a model for how memory retrieval could emerge spontaneously in a simple neural network. This model is almost absurdly simplistic, even stupid in its assumptions. Neurons are either on or off, ignoring subtleties of firing rates or even graded activation. Directionality is also ignored; links between neurons are equally strong in each direction. Exactly how the network is presumed to first arrive in its initial state is left a mystery. Yet analysis of the model showed that something *like* biological neural networks could produce content-addressable memory. Hopfield himself later showed that the model's functioning was robust to the relaxation of some of his strict assumptions (Hopfield, 1984), and the work has laid the foundation for much subsequent work in understanding the neurobiology of memory.

Bass's Model of the Diffusion of Innovations

How do new products diffuse in a population? In the early 1960s, Everett Rogers (1962) provided a near-exhaustive study of this question. He showed that cumulative adoption very often corresponds to an S-shaped curve in which adoption starts slowly, accelerates, and then plateaus. Although Rogers showed that this pattern of product diffusion is common to a startlingly wide variety of domains, he could not explain it. Instead, he merely identified five tautological categories of adopters, defined in terms of their timing of adoption. This explanation is rather unsatisfying and raises many additional questions, including why individuals would fall into a particular category of adopter and how robust the adoption curves are to different proportions of each of those categories. Shortly after Rogers' book was published, Bass (1969) introduced a simple model that provided a strikingly parsimonious explanation of Rogers' data. Suppose, said Bass, that instead of five discrete types, there is only a single type of individual, who with some small probability spontaneously adopts the new innovation (i.e., becomes an innovator) and otherwise adopts with a probability proportional to the number of other adopters he or she encounters. In other words, suppose that innovations spread like diseases. Bass constructed a mathematical model based on these assumptions, and showed not only that they resulted in S-shaped adoption curves, but that by fitting the model to empirical data on the diffusion of different products, characteristics of a given population concerning the rate of observation and the propensity to adopt could be inferred. The Bass model is still the core model for studying the diffusion of products used in communication, technology, and marketing research today (Bass 2004). The Bass model is, of course, extremely simplistic. It ignores real differences between individuals, such as network position (Valente, 1996) or the propensity to adopt based on social group membership (Berger & Heath, 2008), which may influence the dynamics of diffusion. Nevertheless, the Bass model provides critical structure for developing theory and guiding data collection related to the diffusion of innovations.

The Dynamics of Distinctiveness

Some of my own work has concerned the population dynamics resulting from individual preferences for distinctiveness. Though much of human social behavior stems from conformity—that perfectly reasonable heuristic to copy others "when in Rome"—it is also quite common to actively differentiate ourselves from others (at least in the large, complex societies in which most of us find ourselves; see Smaldino in press). I first became involved in this research in graduate school, when I was approached by two social psychologists working within the domain of optimal distinctiveness theory (ODT; Brewer, 1991; Leonardelli et al., 2010). This theory has long had at its core the sort of vague verbal model I discussed in the subsection "Building Models, Formal and Otherwise." The presumption is that individuals have traits called social identities, and that, all else being equal, they will "identify" with whichever identity optimally balances the opposing needs for assimilation (to be similar to others) and differentiation (to be different from others). It is never stated exactly what does or does not constitute a social identity, what it means to identify as one thing, how the needs for assimilation and differentiation are calibrated, or how one optimizes a balance between them. Empirical tests have shown that US college students do prefer to express, at least on paper, a more exclusive part of their social identity when the initially proposed identity (e.g., being a student of their college) is described as being non-noteworthy (Brewer & Pickett, 2002). However, many questions remain, and the theory remains largely lacking in precision.

One assumption of the ODT is that deviations from optimality will be corrected as individuals change their expressed identities to ones that more optimally balance their opposing needs, and that this will result in a stable equilibrium in which individuals are satisfied in their relative distinctiveness (Leonardelli et al., 2010). To test this, my colleagues and I decided to model a simple scenario based on one possible interpretation of ODT (Smaldino et al., 2012). We assumed a population of individuals who could each express one of some number of discrete identities at any given time. We also assumed that each individual had a preference for some optimal level of distinctiveness, where an individual's distinctiveness was defined as the proportion of neighbors also expressing the same identity. One at a time, agents would consider the distinctiveness of their currently expressed identity, and if a better option was available, switch to that identity (agents were updated one at a time because synchronous updating is unrealistic, eliminates the possibility of behavioral cascades, and can generate peculiar model artifacts; see Huberman & Glance, 1993). The result was that individuals *always* ended up expressing identities that were far too popular to satisfy their preferences for differentiation. I later learned that this result echoed earlier work by ecologists considering animals joining groups of varying size, who had reached similar conclusions (Sibly, 1983).

Our model makes extremely simplistic assumptions about individuals' abilities to observe, express, and change identities. Nevertheless, the model accomplishes something that no previous work on ODT had: it defined all of the parts of the system and their relationships *explicitly*. Based on a set of assumptions that is entirely consistent with the verbal model, we produced a model that provided several initial conclusions and prompted two broad questions. First, is it true that individuals are perpetually more similar to others than they would prefer? This could, in fact, be the case. Several other models have recently shown that even explicit preferences for anti-conformity or distinctiveness can nevertheless result in local conformity (Muldoon, Smith, & Weisberg, 2012; Touboul, 2014; Smaldino & Epstein, 2015). Second, if it is instead the case that individuals *are* generally satisfied with the distinctiveness of their expressed identity, then what key factors related to the dynamics of identity expression were missing from our model? Several possibilities present themselves, including factors such as network structure, interdependence between identities, behavioral inertia, and transaction costs to switching identities. We examined the first of these, network structure, by situating individuals on a square lattice and having them only respond to nearby neighbors. We found that for a wide range of conditions, this kind of network structure solved the problem: individuals could maintain identities that maximized their preferences for distinctiveness. Our implementation of network structure was itself quite unrealistic—real social networks rarely approximate square lattices. Nevertheless, the model represents a step, if only a small one, toward a more precise theory linking individual preferences for distinctiveness with the social organization that results from those preferences.

Turning the Modeling Lens on the Scientific Process

As a final example, I want to explore how formal models can help us better understand the larger endeavor in which we are engaged: science. Recently, controversy has raged over the roles of replication and publication policy in improving the reliability of research (Open Science Collaboration, 2015). Some propose that all results should be published, to ensure that a "file drawer effect" doesn't lead to over-representation of positive results (Franco, Malhotra, & Simonovits, 2014), while others are skeptical of the value of failed replications because replication studies may have diminished power (Kahneman, 2011; Bissell, 2013; Schnall, 2014). All acknowledge the importance of replication, but opinions vary widely on how much is needed and what its evidential value might be. Until now, each view has been based on intuition and lacked concrete rationale. And empirical analysis is inherently limited, both by the incompleteness of the published record and by the lack of internally consistent models of the scientific process that would allow us to usefully interpret extant data.

To remedy this dearth, Richard McElreath and I developed an analytical model of the population dynamics of science (McElreath & Smaldino, 2015). The model represents a population of scientists who, with regularity, select a hypothesis for investigation, investigate it using the standard methods of their field, and then attempt to communicate their results to the scientific community. We built on previous work by Ioannidis (2005), who introduced a simple model of scientific investigation that highlights the importance of the base rate—that is, the a priori probability that a novel hypothesis is true. When the base rate is low, even the most stringent experimental methods may produce more false positives than true positives. Our model extends this discussion to consider the fact that scientists may replicate their own and each other's work, but also that results must also run the gauntlet of peer review, with negative results being less likely to be published than positive ones. We conclude that regardless of how much replication is done, the biggest impediments to the effectiveness of science are low base rate and high false positive rate. I know of no better way to improve the base rate than to make sure that hypotheses stem from well-validated, precise theories. Such theories, in turn, are often developed at least partly through the extensive use of formal modeling. The model also speaks directly to the debate over the meaning of failed replications. We show that replications are informative even when they have substantially lower power than the initial investigations. Perhaps counterintuitively, we also find that suppression of negative findings may be beneficial, at least when such findings are tests of novel hypotheses and the base rate is low. Under those conditions, most novel results will be correct rejections of incorrect hypotheses. As these will not be surprising, we may want to avoid filling our journals with such results, or at least delegate them to a distinct location.

Our model of science is extremely simple. It frames hypothesis testing in a standard but unsatisfying true/false classification, rather than considering practical significance and effect size estimation. It ignores researcher bias, multiple testing, and data snooping. It ignores the incentives that drive scientists in choosing and publishing results, as well as differences in exclusivity and impact between journals. Nevertheless, our model provides, for the first time, specific quantitative evaluations of many verbal arguments. As I have noted throughout this chapter, all models, whether formal or verbal, ignore some factors. The difference is that, with a formal model, it is precisely clear which factors are being considered and which are being excluded.

Modelers Are Stupid (Sometimes)

Models can help us to specify theories of how a complex system works, and to assess the conclusions of our assumptions when they are made precisely. However, I want to be careful not to elevate modelers above those scientists who employ other methods. This is important for at least two reasons, the first and foremost of which is that science absolutely requires empirical data. Those data are often painstaking to collect, requiring clever, meticulous, and occasionally tedious labor. There is a certain kind of laziness inherent in the professional modeler, who builds entire worlds from his or her desk using only pen, paper, and computer. Relatedly, many scientists are truly fantastic communicators, and present extremely clear theories that advance scientific understanding without a formal model in sight. Charles Darwin, to give an extreme example, laid almost all the foundations of modern evolutionary biology without writing down a single equation. That said, evolutionary biology would surely have stagnated without the help of formal modeling. Consider that Darwinism was presumed to be in opposition with Mendelian genetics until modelers such as R. A. Fisher and Sewall Wright showed that the two theories were actually compatible.

The second reason is that having a model is not the same thing as having a *good* model, or a model that is well presented, well analyzed, or well situated in its field. I want to focus on presentation and analysis. A model's strength stems from its precision. I have come across too many modeling papers in which the model—that is, the parts, all their components, the relationships between them, and mechanisms for change—is not clearly expressed. This is most common with computational models (such as agent-based models), which can be quite complicated, but also exists in cases of purely mathematical models. I am not a big fan of standardized protocols for model descriptions, as the population of all models is too varied and idiosyncratic to fit into a one-size-fits-all box. I will simply ask modelers to make an effort in their reporting. Make sure your model description is clear. The broad strokes, which may stem from verbal theory, should come first, followed by a filling in of details. When possible, make code available as soon as your paper is published, if not before. Clarity reveals how well the model really represents the system it purports to represent. Obfuscation is the refuge of the poor or insecure modeler.

This is not the place to go into great detail about the best practices for model analysis. I will only say that a major benefit of a model is the ability to ask all manner of "what if" questions. The assumptions of a model, including but not limited to its parameter values, should be explored extensively. After all, obtaining the conclusions that follow from those assumptions is the entire purpose of modeling. If you forgive the indulgence, I'll pick one small nit here concerning methods for analyzing computational models. Where differences between conditions are indicated, avoid the mistake of running statistical analyses as if you were sampling from a larger population. You already have a generating model for your data—it's your model. Statistical analyses on model data often involve modeling your model with a stupider model. Don't do this. Instead, run enough simulations to obtain limiting distributions.

Finally, it is important to always evaluate whether the conclusions of our model rely on reasonable assumptions. For example, it has been claimed that some economists have fallen prey to a sort of theory-induced blindness, giving too much credence to their models—which are generally based on the theory of the rational actor—and ignoring the fact that the core assumptions of the model are based on severe distortions of human psychology (Thaler, 2015). Microeconomic models based on rational choice theory are useful for developing intuition, and may even approximate reality in a few special cases, but the history of behavioral economics shows that standard economic theory has also provided a smorgasbord of null hypotheses to be struck down by empirical observation.

Conclusion

Humans, scientists included, are limited beings who are bad at forming intuitions about the organization and behavior of complex systems. Verbal models, while critical first steps in scientific reasoning, are necessarily imprecise. Overreliance on verbal models can impede precision and, by extension, impede progress in our understanding of complex systems. Formal models are explicit in the assumptions they make about how the parts of a system work and interact, and moreover are explicit in the aspects of reality they omit. This has the potential disadvantage of making formal models appear stupid. And of course, they *are* stupid, because we are limited beings and stupid models are the best we can do. As Braitenberg writes, fiction will always be part of science "as long as our brains are only miniscule fragments of the universe, much too small to hold all the facts of the world but not too idle to speculate about them" (Braitenberg, 1984, p. 1).

An old adage holds that it is better to stay silent and be thought a fool than to speak and remove all doubt. As scientists, our goal is not to save face, but in fact to remove as much doubt as possible. Formal models make their assumptions explicit, and in doing so, we risk exposing our foolishness to the world. This appears to be the price of seeking knowledge. Models are stupid, but perhaps they can help *us* to become smarter. We need more of them.

References

Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, *15*, 215–227.

Bass, F. M. (2004). Comments on "A new product growth for model consumer durables": The Bass model. *Management Science*, *50*, 1833–1840.

Bedau, M. A. (1999). Can unrealistic computer models illuminate theoretical biology? In Wu, A. S. (Ed.), *Proceedings of the 1999 Genetic and Evolutionary Computation Conference Workshop* (pp. 20–23). Orlando, FL: GECC.

Berger, J., & Heath, C. (2008). Who drives divergence? Identity signaling, outgroup dissimilarity, and the abandonment of cultural tastes. *Journal of Personality and Social Psychology*, *95*, 593–607.

Berryman, A. A. (1992). The origins and evolution of predator–prey theory. *Ecology*, *73*, 1530–1535.

Bissell, M. (2013). Reproducibility: The risks of the replication drive. *Nature*, *503*, 333–334.

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge, MA: MIT Press.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*(5), 475–482.

Brewer, M. B., & Pickett, C. L. (2002). The social self and group identification: Inclusion and distinctiveness motives in interpersonal and collective identities. In J. Forgas & K. Williams (Eds.), *The social self: Cognitive, interpersonal, and intergroup perspectives*. Philadelphia, PA: Psychology Press.

Campbell, D. T. (1974). "Downward causation" in hierarchically organised biological systems. In F. Ayala & T. Dobzhansky (Eds.), *Studies in the philosophy of biology* (pp. 179–86). Oakland, CA: University of California Press.

Eisenberg, E. M. (1984). Ambiguity as strategy in organizational communication. *Communication Monographs*, *51*(3), 227–242.

Epstein, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation*, *11*(4), 12.

Farrell, S., & Lewandowsky, S. (2010). Computational models as aids to better reasoning in psychology. *Current Directions in Psychological Science*, *19*: 329–335.

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*, 1502–1505.

Gigerenzer, G. (1998). Surrogates for theories. *Theory & Psychology*, *8*(2), 195–204.

Gleick, J. (2004). *Isaac Newton*. New York: Vintage Books.

Gunawardena, J. (2014). Models in biology: "Accurate descriptions of our pathetic thinking." *BMC Biology*, *12*, 29.

Hills, T. T., Jones, M. N., & Todd, P.M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*, 431–440.

Huberman, B. A., & Glance, N. S. (1993). Evolutionary games and computer simulations. *Proceedings of the National Academy of Sciences USA*, *90*, 7716–7718

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA*, *79*, 2554–2558.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences USA*, *81*, 3088–3092.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124.

Kahneman, D. (2011). A new etiquette for replication. *Social Psychology*, *45*, 310–311.

Kauffman, S. A. (1971). Articulation of parts explanation in biology and the rational search for them. In R. C. Buck & R. S. Cohen (Eds.), *PSA 1970* (pp. 257–72). Irvine, CA: Philosophy of Science Association.

Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative Science Quarterly*, *52*, 667–694.

Leonardelli, G. L., Pickett, C. L., & Brewer, M. B. (2010). Optimal distinctiveness theory: A framework for social identity, social cognition, and intergroup relations. In M. P. Zanna & J. M. Olson (Eds.), *Advances in experimental social psychology* (Vol. 43, pp. 66–115). New York: Academic Press.

Luckinbill, L. S. (1973). Coexistence in laboratory populations of *Paramecium aurelia* and its predator *Didinium nasutum*. *Ecology*, *54*, 1320–1327.

Marewski, J. N., & Olsson, H. (2009). Beyond the null ritual: Formal modeling of psychological processes. *Zeitschrift für Psychologie*, *217*, 49–60.

McElreath, R. (2015). *Statistical rethinking: A Bayesian course with R examples*. New York: Chapman & Hall.

McElreath, R., & Smaldino, P.E. (2015). Replication, communication, and the population dynamics of scientific discovery. *PLoS One*, *10*(8), e0136088.

Moussaïd, M., Helbing, D., & Theraulaz, G. (2011). How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences*, *108*, 6884–6888.

Muldoon, R., Smith, T., & Weisberg, M. (2012). Segregation that no one seeks. *Philosophy of Science*, *79*, 38–62.

Nowak, A., Rychwalska, A., & Borkowski, W. (2013). Why simulate? To develop a mental model. *Journal of Artificial Societies and Social Simulation*, *16*(3), 12.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716.

Popper, K. (1963). *Conjectures and refutations*. New York: Routledge.

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

Rogers, E. M. (1962). *Diffusion of innovations*. New York: The Free Press.

Schank, J. C. (2001). Beyond reductionism: Refocusing on the individual with individual-based modeling. *Complexity*, *6*(3), 33–40.

Schank, J. C. (2008). The development of locomotor kinematics in neonatal rats: An agent-based modeling analysis in group and individual contexts. *Journal of Theoretical Biology*, *254*, 826–842.

Schank, J. C., May, C. J., & Joshi, S. S. (2014). Models as scaffold for understanding. In L. R. Caporael, J. R. Griesemer, & W. C. Wimsatt (Eds.), *Developing scaffolds in evolution, culture, and cognition* (pp. 147–167). Cambridge, MA: MIT Press.

Schnall, S. (2014). Clean data: Statistical artefacts wash out replication efforts. *Social Psychology*, *45*(4), 315–320.

Sibly, R. M. (1983). Optimal group size is unstable. *Animal Behaviour*, *31*, 947–948.

Smaldino, P. E. (2016). Not even wrong: Imprecision perpetuates the illusion of understanding at the cost of actual understanding. *Behavioral and Brain Sciences*, *39*, e163.

Smaldino, P. E. (in press). The evolution of the social self: Multidimensionality of social identity solves the coordination problems of a society. In W. C. Wimsatt & A. C. Love (Eds.), *Beyond the meme: Articulating dynamic structures in cultural evolution*. Minneapolis, MN: University of Minnesota Press.

Smaldino, P. E., Calanchini, J., & Pickett, C. L. (2015). Theory development with agent-based models. *Organizational Psychology Review*, *5*(4), 300–317.

Smaldino, P. E., & Epstein, J. M. (2015). Social conformity despite individual preferences for distinctiveness. *Royal Society Open Science*, *2*, 140437.

Smaldino, P. E., Pickett, C., Sherman, J., & Schank, J. (2012). An agent-based model of social identity dynamics. *Journal of Artificial Societies and Social Simulation*, *15*(4), 7.

Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review*, *11*, 87–104.

Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. New York: W. W. Norton.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20*, 410–433.

Touboul, J. (2014). The hipster effect: When anticonformists all look the same. *arXiv*, 1410.8001.

Valente, T. W. (1996). Social network thresholds in the diffusion of innovations. *Social Networks*, *18*, 69–89.

Weinhardt, J. M., & Vancouver, J. B. (2012). Computational models and organizational psychology: Opportunities abound. *Organizational Psychology Review*, *2*(4), 267–292.
Wimsatt, W. C. (1974). Complexity and organization. In K. Schaffner & R. S. Cohen (Eds.), *PSA 1972* (pp. 67–86). Irvine, CA: Philosophy of Science Association.
Wimsatt, W. C. (1987). False models as means to truer theories. In M. H. Nitecki & A. Hoffman (Eds.), *Neutral models in biology* (pp. 23–55). New York: Oxford University Press.