# Working Models and the Synthetic Method:

## Electronic Brains as Mediators Between Neurons and Behavior

*Peter Asaro*

This article examines the construction of electronic brain models in the 1940s as an instance of "working models" in science. It argues that the best way to understand the scientific role of these synthetic brains is through combining aspects of the "models as mediators" approach (Morgan and Morrison, 1999) and the "synthetic method" (Cordeschi, 2002). Taken together these approaches allow a fuller understanding of how working models functioned within the brain sciences of the time. This combined approach to understanding models is applied to an investigation of two electronic brains built in the late 1940s, the Homeostat of W. Ross Ashby, and the Tortoise of W. Grey Walter. It also examines the writings of Ashby, a psychiatrist and leading proponent of the synthetic brain models, and Walter, a brain electro-physiologist, and their ideas on the pragmatic values of such models. I conclude that rather than mere toys or publicity stunts, these electronic brains are best understood by considering the roles they played as mediators between disparate theories of brain function and animal behavior, and their combined metaphorical and material power.

*Keywords:* models, psychology, cybernetics

> It seems to me that this purely pragmatic reason for using a model is fundamental, even if it is less pretentious than some of the more "philosophical" reasons. Take for instance, the idea that the good model has a "deeper" truth– to what does this idea lead us? No electronic model of a cat's brain can possibly be as true as that provided by the brain of another cat; yet of what use is the latter as a model? Its very closeness means that it also presents all the technical features that make the first so difficult. From here on, then, I shall take as a basis the thesis that the first virtue of a model is to be useful. (Ashby, 1972: 96).

In the first half of the 20$^{th}$ century a small scientific movement emerged and took an unexpected turn, beginning an intriguing journey towards a mechanistic understanding of the mind. What made this journey intriguing was that much of it was not focused on humans nor animals nor brain tissues, but on machines–electronic models of the brain. While in some sense a revival of 17th century mechanistic mental philosophy, this era marked a departure from the dominant 19th century mental philosophy of rationalism, as well as the empirical approaches of neurophysiology and psychopathology. Yet, it was these strange electronic brains which eventually proved to be central in forging fundamental connections between those different traditional approaches to the mind, and provided a sense that a comprehensive understanding of the mind, brain and behavior was within the grasp of science. A new found openness to empirical approaches to mental philosophy following pragmatists like William James, and an increasing diversity of new empirical methods coming from psychology and neurology led to an experimental era in the mental sciences which sought an empirical basis for understanding behavior and mental activity. The ensuing cognitive revolution in the brain sciences depended upon these electronic brain models in an essential way–indeed the digital computer was in many ways itself conceived and constructed as such an electronic brain model and eventually served as the central metaphor for the brain. This article aims to understand the nature and role of these machines as working models. It was because these electronic brains were able to function as working models that they were so successful in advancing the mechanistic view of the mind.

While the Cybernetics movement of the 1940s and 1950s is often cited as a precursor to the cognitive revolution in the late 1950s brain sciences, its role is often seen as limited to providing general inspiration and some rudimentary theory. I argue, however, that its role was far more significant, despite the fact that the particular details of its theories were largely discarded by mainstream cognitive science in the following decades. In particular, little attention has been paid until recently to the devices that the cyberneticians actually built, and how they devised a scientific methodology around the construction of their synthetic brains. Roberto Cordeschi (2002) has written a history of the development of these devices beginning in the first years of the 20th century, and has described their approach as the "synthetic method." Though such a methodology had been used in other sciences for centuries, the weaving of constructed technologies and scientific explanation into a synthetic approach was certainly novel to the brain sciences. In his history, two devices stand out as exemplary cases of this methodology: The Homeostat and the Tortoise, designed by the British Cybernetics pioneers W. Ross Ashby and W. Grey Walter, respectively, in 1948. While Cordeschi (2002) presents much of the historical background for these devices, and recognizes the new methodology at work, he does not go so far as to provide an analysis of how the new synthetic method actually functioned, or why it succeeded. This is precisely what I aim to do in this article, and I believe the key to understanding the synthetic method lies in the epistemology of working models and their role as mediators between disciplines and theories.

In recent years researchers in the social studies of science have begun to pay a great deal of attention to scientific models. There are many views of what models are, but they are traditionally associated with abstractions–mathematical or cognitive models–and are often treated as the metaphysical glue which binds theory to empirical data (Hesse, 1966). The tradition in the philosophy of science that has paid the most attention to models has been the anti-realists. Beginning with Bas van Fraassen (1980), those seeking to challenge realist interpretations of scientific theories have turned to an examination of the role of models. While van Fraassen's approach placed a heavy reliance on what he called "partial isomorphisms," similar approaches were followed by Ian Hacking (1983) and Nancy Cartwright (1983) that placed less significance in isomorphism as an epistemic property, instead placing it in scientific practices more generally. For Hacking these practices are representing and intervening, while for Cartwright they involve capturing the phenomenal qualities of natural events, and the causal structure of those events when they can be found. Ronald Giere (1999) adopts a very similar account of models, but in the context of a somewhat different project–that of offering a cognitive understanding of science. Finally, the notion of an ideal or complete isomorphism, or what Paul Teller (2001) has called the "Perfect Model Model," has been identified as underlying much of the philosophical theorizing of models. He demonstrates that the perfect model model is untenable and frustrates our attempts to get a clear picture of the role of models in science.

In this article, I will not seek to supplant the abstract view of models, which I will call "theoretical models," but rather will seek to add to it the concept of a different kind of model which also appears in science–the "working model." Working models differ from theoretical models in that they have a material realization, and are thus subject to manipulation and interactive experimentation. It is this dynamic material agency which sets working models apart from other closely related kinds of scientific objects, such as theoretical models, simulations and experimental instruments.

This is not to say that they lack any theoretical basis or significance, indeed their relevance to the broader field of scientific knowledge depends upon having some form of theoretical implication. However, unlike theoretical models they are not solely abstract constructs, nor merely the practical formulation of a theory for some specific application. By having a material face, working models participate in the world of material culture. In virtue of this, working models function in significantly different ways than theoretical models function. Most importantly, working models exhibit what Pickering (1995) calls material agency–in virtue of their materiality they behave in ways that are unintentional, undesired and unexpected to their designers. This is by no means a negative feature, and it is often essential to the productive contribution of such models– such as becoming oracles, providing new insights, presenting novel phenomena, and offering serendipitous knowledge. The dynamic material agency inherent in working models is prerequisite for autonomous and open-ended interactions with the environment and with experimenters–interactions that emerge temporally. It is in the dynamic flow of interactions between researchers and material models that new techniques are developed, and practical knowledge of the system and its behavior emerges. Accordingly, through these interactions scientific understanding can be extended–new explanations can be derived and tested, and new phenomena can be observed, manipulated and brought under control.

While I wish to single out working models for more careful scrutiny, I do not necessarily wish to argue for a strict distinction between them and other closely related scientific objects–such as simulations, instruments and experimental apparatus that often share many of the same properties and participate in the same material culture of science. In the universe of scientific models, working models fall somewhere between simulations and instruments, or might even be seen as a sort of hybrid of the two. Clearly defining the boundaries between these, if such a project is possible, is well beyond the scope of the present article. It is however useful to note the similarities to other kinds of models described in the science studies literature, in particular discussions of simulations, scale models, and models as mediators can all help inform an understanding of working models. I will review some of these briefly before examining the case of working models in the brain sciences of the 1940s.

## Models in Science Studies

> From this point of view, there is no such thing as the true model of such a complex system as a cat's brain. Consider, for instance, the following four possible models, each justifiable in its own context:
>
> 1. An exact anatomical model in wax.
>
> 2. A suitably shaped jelly that vibrates, when concussed, with just the same waves as occur in the real brain.
>
> 3. A biochemical soup that reacts biochemically just as does the cat's brain when drugs are added.
>
> 4. A programmed computer that gives just the same responses to auditory stimuli as does the cat's brain.
>
> Clearly, complex systems are capable of providing a great variety of models, with no one able to claim absolute authority. (Ashby, 1972: 97)

Recently there has been a growing interest in the nature of computational simulations and their use in science. In one of the more carefully thought out analyses, Winsberg (2003) examines the use of simulated experiments in physics. He identifies these simulations as a scientific practice that lies somewhere between traditional theorizing and traditional experimentation. In these simulations, theory is applied to virtual systems in an effort to test and extend the theory. Like experiments, simulations have what Hacking (1992) calls a "life of their own." Hacking originally applied this notion to thought experiments, but it implies that there is an autonomous agency to the simulated experiments – they are not simply an expression or extension of the experimenters. For Winsberg, this autonomous agency expresses itself through the life-cycle of a simulated experiment – the recycling and retooling that goes into developing a useful simulation responds to the autonomous agencies expressed in the practices of simulation over time.

Working models might easily be construed as simulated experiments in this sense. That is, they too lie somewhere between theory and experiment in the traditional senses, and have a “life of their own.” However, it is important to acknowledge the material basis of these simulations. While it may be tempting to conceive of computational simulations as virtual systems devoid of materiality, this is not really true. Computers are very sophisticated material artifacts, and their calculations are not ethereal or disembodied. A vast number of circuits and electrical currents are involved in realizing a computational simulation. But the nature of computational simulations, which are really an outgrowth of pencil-and-paper mathematical models, is to realize a mathematical formalism as precisely as possible. While certain mathematical techniques can never be realized by computers, the general aim of these simulations is to at least approximate the proper formalism. As such, the materiality of these computational simulations is usually disparaged as the cause of errors, rather than as a source of any insights into the theory. Working models, by contrast, are more likely to embrace their materiality, rather than hide it. While they often start by realizing a mathematical theory, they further aim to demonstrate phenomena that may not be easily expressed by mathematical formalisms. And the “life” of the working model dwells heavily in the material realm, though it also gets expressed in the iterative redesign of models.

Recent work on scale models in architecture, especially Yaneva (2005), demonstrate that these models function in many ways similar to the working models I will describe here, and embrace their material nature. These models are used in architectural practice to develop a subjective understanding of interior and exterior space that is not possible in drawings or other 2-dimensional representations. These scale models of buildings are involved in an iterative design process that is autonomous from certain engineering concerns and constraints, but which necessarily conforms to spatial requirements. These scale models are thus a kind of working model, though a somewhat limited form due to their static character. That is to say that the working models I find most interesting are the dynamic models that exhibit various behaviors and emergent properties. While architectural scale models do this, it is only in the sense that their forms evoke a cognitive response in designers and architects, who largely depend on expertise to “read” these artifacts. They are thus a form of sophisticated image or visualization, and as such their use and significance outside of that practice is limited. Working models, due to their dynamic character, are often found outside of their scientific contexts and are often admired by non-experts, as we shall see with the Homeostat and Tortoise.

Perhaps the most insightful perspective on models in the recent science studies literature has been on their role as mediators. In the introduction to their edited collection on the subject, Morgan and Morrison (1999) articulated the notion of models as mediators:

> [W]e want to outline... an account of models as *autonomous agents*, and to show how they function as *instruments* of investigation... It is precisely because models are partially independent of

> both theories and the world that they have this autonomous component and so can be used as instruments in the exploration in both domains. (Morgan and Morrison, 1999: 10)

The idea that models are autonomous agents is crucial to a pragmatic understanding of science and its material culture. In this statement we can see that the notion of models they have in mind is one which stands in between theory and the world–whether "the world" is construed as phenomena, data or experiment. And we see in the various cases described in the volume that models are seen as offering a space of play in a conceptual realm where theory and world are not strictly related, but their inter-relation is in play through the mediation of models. While this is a nice image, and captures important aspects of the presented cases well, it is not the only notion of autonomous agency operative in the working models of the brain sciences. There, the mediation is being done between multiple theories – the electronic brains mediated between theories of neurons and theories of behavior, not between theory and data – and the agency also lies within the material world.

Moreover, Morgan and Morrison's notion of autonomy is more concerned with the fact that models are *not determined* by theory or by data–that they have freedom from both. But all this comes to in practice is that scientists are able to manipulate models freely, where theories are constrained formally and data are constrained empirically. Models then become a plastic medium (at least to the extent that it is not constrained formally or materially) and is subject to willful manipulation by human agents. It is difficult to find much agency in the models themselves in these accounts, though the models do become the focus of much scientific work.

I am not the first to recognize the significance of working models in the brain sciences. Not only did the scientists who devised those electronic brains engage in a critical reflection on the nature and use of their own models, but Cordeschi (2002) has written a wonderful history of the development of robots and mechanistic psychology up to and including Ashby and Walter. In his account, the synthetic method is supposed to stand against an analytic method–according to which understanding is obtained by isolating and manipulating various factors in a phenomenon until one can determine the control variables of the phenomenon and their causal effects. According to the synthetic method, scientific understanding can also be derived through constructing a complex model, and examining its properties and behaviors. There is, of course, an element of analysis involved in the determination of which aspects of the phenomena to synthesize in one's model, but the scientific practices of experimentation are instead focused on the synthetic model, rather than the original phenomenon. The cyberneticians were also interested in the formal sense of this distinction, seeing the brain as a multi-variable system of such great complexity that analytic mathematical techniques would fail to reveal its secrets, and so it required radical new techniques. The computational models of complex systems developed by John von Neumann, and his use of Monte Carlo methods in particular, are one of the clearer examples of this

(Aspray, 1990). In this article, I do not wish to stick to this strictly formal definition, however, and prefer to use "synthetic" to connote the constructed and dynamic character of models, rather than the nature of the mathematical techniques they employ.

When viewed in the context of recent science studies, the synthetic method is a methodology that derives its efficacy by enhancing the material practices that scientists can employ in studying a phenomena. This allows science to circumvent theoretical deadlocks, as well as begin the investigation of phenomena which are not otherwise accessible experimentally.

I believe that the "synthetic method" and "models as mediators" approaches can be brought together to inform our conception of working models. It would seem that a complete explanation of the epistemic basis of the synthetic method requires an explanation of the scientific context in which such models are constructed. At least, this is the case if we wish to explain the construction of the synthetic brain models as contributions to the understanding of cognitive neuroscience in the 1940s. And in this case, we can only understand the role of these models by examining the mediations they achieved between the disjoint and incomplete theories being developed in the various disciplines studying different aspects of the mind, brain, and behavior.

It is important to recognize that there are multiple ways in which a model can serve as a mediator. While the models as mediators literature focuses primarily on the role of models as mediating between theory and data, there are at least two other senses of mediation which I want to focus on in this article. The first is the notion that a model can mediate between two theories. The theories involved could be very similar, or could refer to different entities, different types of entities, or the same phenomena at different levels of analysis. In each case it is sometimes possible to build a model which includes key aspects of those theories, and shows how they might work in conjunction, reinforce one another, or even suggest that a theoretical "synthesis" of the two is possible. The second is that models, including the simulations and scale models just described and working models in particular, have the ability to "stand in" for natural systems and phenomena during experimental investigations. This ability is an integral part of the synthetic method that Cordeschi describes. By building the synthetic brains according to certain *principles of construction*, these models were convincingly argued to exhibit certain theories of mind and mechanisms of behavior. By mediating between theories and simultaneously acting as a stand-in, these models provided an innovative approach to the sciences of the mind.

## The Tortoise and the Homeostat

> In order to study this [feedback] abstraction more easily, models have been built containing only two elements connected with two receptors, one for light and one for touch, and two effectors giving progress and rotation, with various possibilities of interconnection. This device is in the nature of a toy rather than a tool and reminds one of the speculations of Craik and the homeostat of Ashby. (Walter, 1953: 3).

W. Ross Ashby began designing the Homeostat in 1946, and gave his first public demonstration at a meeting of the Electroencephalorgaphy (EEG) Society in 1948. His design actually began with a set of mathematical functions which exhibited a property Ashby called *ultra-stability* (similar to what is now called *convergence*), meaning that iterated computations would eventually result in stable, unchanging values or a repeating sequence. He then set about to devise a mechanical or electronic device that would exhibit this behavior. After various attempts he arrived at an electro-mechanical design that included four Homeostat units, each receiving an input current from each of the other three units and sending an output current to each of the other three units (Ashby, 1952). The units themselves were built out of war surplus parts, mainly old radar units. Each unit consisted of a black box with four rows of switches, and a water trough on top with a movable needle resting in the water. The state of the machine is displayed by this needle, and the machine is “stable” when the needle is in the center of the trough, and “unstable” when the needle moves to one end or the other, or moves back and forth. The currents from the other units are routed through a resistor, selected from a fixed set of resistors by a dial, to the water trough, where the needle responds by moving in relation to the current gradient created in the trough.

The behavior of each unit depends upon its specific configuration of switches and dials. In the mundane configuration, each unit merely displays the summation of currents by the position of its needle, and the needle is only stable, *i.e.* in the middle of the trough, by coincidence. However, when the unit is switched so that, instead of a simple resistor, its Uniselector circuit is engaged, it becomes ultrastable. In this configuration, when the needle is not in the center of the trough, a circuit is closed which charges up a coil (capacitor) that, when it passes a certain threshold, discharges to the Uniselector causing it to change to a new resistance (chosen from a pre-arranged randomized set of resistance values). It is thus a sort of random resistor and instantiates a trial-and-error search to find a resistance that stabilizes the needle. The result is that the unit will continue changing its internal organization of resistance values until it finds an equilibrium with the needle in the middle of the trough.

Because the four units are interconnected, they must each find an equilibrium in the environment of inputs supplied by the other units, in other words they must all reach an equilibrium at the same time. If a particular state is unstable, a slight disturbance–like pushing a needle out of place–will throw the needles out of equilibrium and the unstable units will continue searching until they find another equilibrium state. Any number of additional disturbances can be introduced, including switching a resistance value or an input’s polarity, holding a needle in place, or even tying two needles together or to a rod that forces them to move in unison. These sort of interventions provide a basis for systematic experimentation with the device that we will consider shortly. Thus, by a mechanism of searching through possible reorganizations by random trial and error, the Homeostat will eventually find its desired equilibrium.

W. Grey Walter’s Tortoise was a different animal than the Homeostat. In fact, Walter was fond of describing his Tortoise as an electronic animal, while the Homeostat was described as an electronic vegetable. Still the devices were built in the same year, 1948, and exhibited a similar regard for the importance of feedback. The Tortoise was really a simple autonomous robot, based on a 3-wheeled chassis built from war surplus radar parts and the mechanical gears from an old gas meter. The two rear wheels spun freely, while the front wheel was driven by an electric motor, and could rotate back and forth or all the way around a 360$^{o}$ rotation by a second motor. Normally, both motors moved at a constant speed and direction, and the resulting motion of the robot was to travel in an elliptical sort of path, at least until it was disturbed. The two motors were subject to the combined control of two sensors. The first sensor was a simple contact switch between the tortoise-like shell of the robot and its base, such that any collision or contact with an object or obstacle would push the shell against the base and close the circuit. The Tortoise would react to this as a collision by reversing its drive motor. The other sensor was a photocell set atop the front wheel assembly, such that it always pointed in the same direction as the front wheel as it rotated about. As such, it acted as a scanning device, and reacted to bright lights. When the photocell received enough light, it would temporarily stop the motor that caused the turning (but not the drive motor) resulting in the robot driving in a straight path towards the bright light.

## Working Models and the Synthetic Method

Roberto Cordeschi's (2002) brilliant history of what he calls the *discovery of the artificial* tells the story of the building of behavioral models before, during and after the cybernetic era in of the 1940s. Central to this story is the *synthetic method* and how it aided the restoration of psychology as a legitimate science through the construction of electronic, robotic and computer simulations of the mind and brain. But how were these synthetic brain models able to do this? Cordeschi's account of the synthetic method draws upon the concept of a working model, but does little to define or explain it. He does provide one key to understanding working models with the notion of a model's *principles of construction*. While he has described the historical context and development of these models, I wish to consider what those developments can tell us about the epistemic nature and scientific use of models. I thus want to start where Cordeschi leaves off.

Cordeschi frequently stresses the importance of working models over other types of models. Unfortunately, he offers only brief comments on just what makes them so desirable or effective. He does, however, contrast working models with analogies:

> Mechanical analogies for nervous functions, however, occupy a secondary position in the present book, compared to the, albeit naive, *working models* of such functions which were designed or physically realized. Those analogies are discussed in some sections of the present book because examining them allows one to clarify the context in which attempts to build those working

> models were made (Cordeschi, 2002: xvi).

Cordeschi thus finds the physicality of working models to be an empirically powerful force. Implicit in this is that actually realizing a model that "works" imposes some significant constraints on free-wheeling theorization. An important aspect of the constraints on building a model is that it must be designed, and the technical achievement of getting a physical model of this kind to work implies that the design is successful, along with the principles that underlie the design. We should also note in the above passage and throughout Cordeschi's book that he does not consider any hybrids which might lie somewhere in between working physical models and descriptive analogies. However, this is precisely where we might wish to place many computer programs and simulations – part material and part formal.

Cordeschi also sees that working models, as simulated experiments, can be used to test hypotheses:

> Working models or functioning artifacts, rather than scanty analogies, are the core of the discovery of the artificial, and in two important ways. First only such models, and not the analogies, can be viewed as tools for testing hypotheses on organism behavior. The way to establish the possibility that complex forms of behavior are not necessarily peculiar to living organisms is "to realize [this possibility] by actual trial," as Hull wrote in 1931 about his own models of learning... The second way in which working models, rather than analogies, are at the heart of the discovery of the artificial [is that] behavioral models can be *tested* (Cordeschi, 2002: xvi).

Though it is not clear to me what the distinction between the two forms of testing is meant to be, there do seem to be two different kinds of demonstration going on. The first point that he makes is really that working models were a demonstrative "sufficiency" argument in the debates over mechanistic approaches to biology going on at the time. This kind of argument for models seeks to demonstrate that mechanisms of certain types are sufficient for certain complex behaviors. By demonstrating that an inorganic mechanism is capable of something assumed to be a strictly organic or biological function, one can falsify an argument to the contrary. This is a fairly weak form of argument, but was crucial at the time to defend the development of an empirical psychology and to take it in a mechanistic direction.

The second point in Cordeschi's discussion of working models is about scientific methodology – that they can be used to actually test theories. This is different from the ways in which theoretical models have been argued to serve in the confirmation of theories, however. Theoretical models are typically argued to serve as bridges between theoretical entities and empirical elements (data or observations). They thus offer explanations of experimental results by relating theoretical explanations to real phenomena. This is the traditional form of mediation performed by models–that between theory and data–and it purposely obscures the work done in configuring material reality and setting up a controlled experiment in order to make the demonstration effective. Working models exercise their agency from the material domain and generate real phenomena which are also in some sense *simulations*. The metaphysics of simulations is rather complicated, especially insofar as one is tempted to argue that there is an easy distinction to be made between "the real" and "the simulation" which has metaphysical import. Of course, a working model (unlike theoretical models) actually *does something* and in virtue of that is part of a concrete dynamic material reality (where theoretical models remain in the realm of abstract static concepts). So synthetic brains may not be real brains, but they are real electronic devices. In virtue of this they generate real phenomena which demand their own explanation. That is to say, the machines exhibit real behaviors, even if they are not real brains.

There are thus two senses of "demonstration" that can operate in working models. The first sense is that of "demonstration proof" in which the working model is brought directly to bear on theoretical debates. The second sense is the communicative and pedagogical sense of a "vivid demonstration." While the basic notion of demonstration is the same, these kinds of models can continue to be effective long after the scientific debates are settled. In this sense, science museums continue to educate the public with displays that feature demonstrations that were at one time crucial to a theoretical debate, and students still replicate experiments, such as Galileo's experiments with acceleration and inclined planes, as a pedagogical device. In fact, the cyberneticians used many of their electronic devices in the classroom as well as the laboratory. We now turn to the cybernetician's own analysis of their use of models, beginning with their demonstrative abilities.

## Working Models as Demonstrations

I now want to explore the account of models and simulations offered by the cyberneticians themselves, and in particular the one offered by W. Ross Ashby. This will bring us back to the specific virtues of working models that Cordeschi is concerned with, as well as the significance of the principles of construction for such models. Ashby gave a great deal of serious thought to how biologically-inspired machines could serve scientific discourse. One of the key values of models, especially working machines, he arrived at was the vivid communication of specific scientific principles:

> **Simulation for vividness.** Simulation may be employed to emphasize and clarify some concept. Statements about machines in the abstract tend to be thin, unconvincing, and not provocative of further thinking. A model that shows a point vividly not only carries conviction, but stimulates the watcher into seeing all sorts of further consequences and developments. No worker in these subjects should deprive himself of the very strong stimulus of seeing a good model carry out some of these activities "before his very eyes." (Ashby, 1962: 461).

As the use of the term "vividness" conveys, concepts are here considered to be like images, either literally or analogically. Much can be said about the metaphorical and analogical strength of theories, and it might be debated whether this compels acceptance of the theory, or aids in its application to new problems, or in fact does only superficial work. Vividness points again to the cognitive component of science, and that the ease of comprehension and understanding is important for the proliferation of knowledge.

The analogical strength of a model is somewhat disparaged by Cordeschi, in favor of the pragmatics of working models, but the value of a good analogy cannot be completely denied. Metaphors, analogies and images have a relational component, the similarity relations that are used to bring out aspects of real phenomena. They also serve as a sort of connective tissue that draws together different ideas and shows their relation–analogy is itself a form of mediation. While much of the epistemic work lies outside the specific relation, the similarity relation serves an important role in organizing various practices–instantiating, demonstrating, measuring, verifying, extending, etc. A central metaphor or analogy, like isomorphism, can be the goal towards which those epistemic practices aim, and either achieve or fail to achieve. Models in this sense can be the compelling exemplars described by Kuhn (1962) around which scientific paradigms are organized. These are simultaneously demonstrations of accepted explanations of some controlled phenomena, and the basis of theoretical and experimental extension through puzzle-solving. They can also be productive metaphors in the way that 18th century anatomical drawing furthered medical understanding through images and their metaphorical relations, as discussed by Barbara Stafford (1991).

Ashby's notion of simulation for vividness is closely related to the notion of demonstration in science. Demonstration devices go back to the beginnings of modern science. Schaffer (1994) has written on the use of mechanical demonstrations of Newtonian physics as being crucial to their adoption in 17th century England. While much of his history focuses on the academic political matrix in which Newtonian physics sought to establish itself, it was the engineers who built the demonstration devices, and the showman-like scientists who demonstrated them that thrust Newton's mechanics onto the English scientific community of the time. Indeed the Homeostat and Tortoise were often used by their builders to promote the concepts and promise of cybernetics to the public at large, and to fellow scientists. As a means to popularize the emerging science of cybernetics, the Tortoises were hugely successful, appearing at the Festival of Britain, on BBC television and several *Life* magazine articles from 1950-1952.

While demonstrations often aim to convince skeptical members of the scientific community of some theoretical understanding, there is also a strong pedagogical function inherent in models when they are utilized in the education of young scientists and engineers. Their ability to convey ideas vividly applies not only to other members of the scientific community, but also to those being introduced to a field as students, and the wider public. By having students observe, or even better, *build* such devices, they come to appreciate the power of simple feedback mechanisms. Thus, the communicative power of working models was to be used not merely to convey new ideas to the existing scientific community, but to train and inspire a new generation of scientific researchers through the transfer of scientific and engineering *practices*. Indeed, as Pickering (forthcoming) points out, in the later part of his career Ashby saw his role as the elder statesmen of cybernetics to be that of education. Toward that end he built several devices intended purely to demonstrate cybernetic principles in the classroom. The pedagogical advantage of a working model, as opposed to a theoretical model, is that students actually learn laboratory techniques by building such a model themselves, and can interact with the finished model in a multi-modal way. That is, while some students learn best through text and some through equations, some may learn best through visual forms, others through tactile forms. Hands-on experience conveys practical and often tacit knowledge in ways that formal expressions of knowledge can rarely achieve. Moreover, the "hands-on" experience develops laboratory skills and intuitions that are crucial to scientific practice.

The Homeostat was, as its name implied, intended to demonstrate the biological principle of homeostasis. According to this principle, an organism would adjust its condition through whatever means were accessible in order to maintain certain critical conditions within the organism. If an animal was cold, it would seek out a warmer place, if its blood pressure got too low, it might increase its heart rate, etc. Ashby's insight was that a general purpose learning mechanism could be derived from this principle. If an organism or machine were allowed to search randomly through its possible actions in the world, it could find for itself ways to maintain the critical conditions of its internal environment. This would work even when the system gets disturbed by unexpected outside influences, including malicious experimenters. The bottom line is that a mechanism with a feedback-induced random search could exhibit the same principle of homeostasis that living creatures did. It was thus an adaptive system, and demonstrated Ashby's extension of it to the concept of ultrastability.

Indeed, in reading *Design for A Brain* (Ashby, 1952) it is often difficult to distinguish Ashby's empirical arguments from his pedagogical illustrations. He moves so rapidly between arguments in favor of the homeostatic principle as being an explanation of certain aspects of the brain's behavior, to analogies between the behavior of the Homeostat and various brain phenomena, to arguments that the Homeostat embodies the principle of homeostasis, that these all seem to hang together. It is difficult to say whether this style of rhetoric is a contribution to the synthetic method or a product of it, as the distinctions between brain, theory and working model all begin to blur into systems based on a shared set of underlying principles.

Unlike Ashby's Homeostat, which was built on an explicit set of formal equations, W. Grey Walter's robotic Tortoises were built upon a rather loose set of principles of construction, and his work on these models correspondingly focused more on the phenomena that they generated than on the demonstration of any specific formal principles. The general principle that was demonstrated was simply that a small number of feedback loops, two actually, could generate a large number of behaviors through their interactions with one another and the environment. Walter frequently argued that much of their value lay in the fact that it was a *demonstration proof* that simple mechanisms could exhibit complex biological phenomena. But it was the character and range of behaviors that he focused on.

In particular, he argued that the various behaviors of his robots were most easily described using biological and psychological terminology. The biological principles that he claimed could be found in these various behaviors included: parsimony (they had simple circuits, yet exhibited numerous and apparently complex behaviors), speculation (they explored the world autonomously), positive and negative tropisms (they were variously attracted and repulsed by lights of different intensities, depending on their internal states), discernment (the machine could sublimate long-term goals in order to achieve short-term goals like avoiding obstacles), optima (rather than sit motionless between two equally attractive stimuli, like Buridan's ass, they would automatically seek out one stimulus, and then perhaps the other), self-recognition (by sensing and reacting to its own light in a mirror), mutual recognition (by sensing the light of another Tortoise), and internal stability (by returning to their hutch to recharge when their batteries go low) (Walter, 1953). Rhodri Hayward (2001) has also shown how Walter used the Tortoises to argue for the fundamental simplicity underlying all human behavior and especially emotion and love.

The Homeostat was a demonstration proof of the power of the homeostatic principle to explain brain phenomena. The principle of homeostasis was already well known in biology, and feedback controllers were widely studied by electrical and mechanical engineers. What was new to the Homeostat was the notion that a set of four interconnected feedback controllers could not only be stable, but would always tend towards stability if each unit were allowed the additional capability of randomly changing its relation to the others. Marvin Minsky (1961) would later credit Ashby with originating the concept of random search that was exploited in much of AI research. His device coupled two key theoretical concepts: behavioral adaptation, and trial-and-error search, and thereby served as a mediator between psychological and computational theories of the brain and behavior. While not explaining any new data, or relating data to theory, this model actually mediated between theories in two disciplines, and at two levels of abstraction. In doing this, it provided a conceptual bridge between the disciplines.

But did this *require* a working model? From a purely theoretical perspective, the Homeostat proved nothing that could not be shown mathematically on paper. There have subsequently been formal proofs for the convergence (or lack of convergence) of a great many search algorithms and random search strategies. In fact, Ashby developed the mathematical equations defining the Homeostat before he began designing the device to realize them, and spent several years attempting various schemes for constructing a device which would both conform to these equations and provide a vivid demonstration of its underlying principles. But unlike a purely mathematical demonstration, the Homeostat made it vividly explicit just how such a random search could be coupled to behavior in an explanatory way.

And so the initial interest and most striking aspect of these working models of the brain was their very existence and demonstration of the fundamental principles of mechanistic psychology. Moreover, the real power of these models lay in their ability to be subjected to experimentation. This was not only more efficient than mathematical analysis at a time when the first digital computers were still under construction. Eventually, the increasing power, ease of use and accessibility of computers would make digital simulations far more attractive than the construction of analog machines for many purposes. But it was the electronic brains that paved the way to computational simulations.

## Working Models as Experiments

> The outlook for an experimental model brightened at once with the problem reduced to the behaviour of two or three elements. Instead of dreaming about an impossible 'monster,' some elementary experience of the actual working of two or three brain units might be gained by constructing a working model in those very limited but attainable proportions. (Walter, 1953: 109).

Now that we have considered some of the demonstrative functions of models, it is time to consider the more traditional roles of models in scientific explanation and verification, and what working models can offer here that theoretical models cannot. It is these aspects which Cordeschi suggests are the specific advantage of working models over "mere metaphors." Ashby characterized this function in terms of deduction and exploration:

> **Simulation for Deduction and Exploration.** Perhaps the most compelling reason for making models, whether in hardware or by computation, is that in this way the actual performance of a proposed mechanism can be established beyond dispute... Note, for instance, the idea that the molar functioning of the nervous system might be explained if every passage of a nervous impulse across a synapse left it increasingly ready to transmit a subsequent impulse. Such a property at the synapse must impose many striking properties on the organism's behavior as a whole, yet for fifty years no opinion could be given on its validity, for no one could deduce how such a system would behave if the process went on for a long time; the terminal behavior of such a system could only be guessed. (Ashby, 1962: 463).

There are several ideas packed into this passage that bear commenting upon. The first thing to note about this passage is that it points to how the availability of a specific set of material practices influences what can be tested, and thus what can become knowledge. Even though some of the mathematical techniques for testing various hypotheses might be available, the work required to carry them out often was too great to be realized. The more important aspect is that working models support the empirical testing of certain hypotheses. Even though the applicability of the hypothesis to the brain might be tenuous, it is still possible to make progress in the details of theory using such models.

As one popularizer of cybernetics put it:

> When we have thus constructed a model which seems to copy reality, we may hope to extract from it by calculation certain implications that can be factually verified. If the verification proves satisfactory, we are entitled to claim to have got nearer to reality and perhaps sometimes to have explained it. (de Latil, 1957: 225).

Whether this constitutes explanation or not might be stretching things, but there are certainly characteristics of a working model that enliven speculation into both the structure and organization of the natural phenomena, and into possible extensions to the working model. I believe that it is the ability to move back and forth between deductions from theories and extensions to models that makes the synthetic method so productive. Though it is in some sense dependent on antecedent and subsequent analytic methods and theorizing, it greatly enhances these by offering new ideas and directions for extension when theorizing and analysis are at a loss. This also helps to explain why it is often difficult to separate the scientific from the technological advances made by the synthetic method – the two become intertwined as technoscience.

There are two senses in which a model can be extended. The first sense does not necessary involve any changes to the model itself:

> Once the model has been made, the work of the model-maker has reached a temporary completeness, but usually he then immediately wishes to see whether the model's range of application may be extended. The process of extension... will be subject to just the same postulate as the other processes of selection; for, of all possible ways of extending, the model-maker naturally wants to select those that have some special property of relevance. Thus, a model of the brain in gelatin, that vibrates just like the brain under concussion, is hardly likely to be worth extension in the biochemical direction. From this point of view the process of extension is essentially an exploration. So far as the worker does not *know* the validity of the extension, to that degree must he explore *without* guidance, *i.e.*, "at random." (Ashby, 1972: 110).

Extension can consist of two different processes: validation (or verification) and exploration. In the first sense, a new model is extended by showing it models a wider range of phenomena than first believed, or it is modified to achieve this effect. In the second sense, a model can be explored to find new kinds of phenomena in it, or built upon and altered to produce new or clearer phenomena. In exploration, one does not seek out some specific correspondence between model and world, but rather one treats the model as itself the object of investigation and extension.

If a model actually generates a genuine behavior in the world, that is a phenomenon itself to be explained. It is not, in this sense a prediction about what might happen, it is itself a happening. Moreover, as with experimental apparatus, phenomena can be played with, changed slightly without any particular expectations as to what new phenomena may result, just to see what might happen. In this sense it is not predictive, nor does it necessarily attempt to fit the data obtained from some other phenomenon. It does not tell you what you might find in the world, it *is* what you find. There is a connotation to the use of the word "experiment" which means open-ended exploration of this sort – a "playing around" or fiddling with things just to see what might happen (Pickering, 1995). A material model like the Homeostat can be played with by students and novices, as well as experts to reveal all sorts of interesting phenomena. Hypotheses about its behavior can also be rigorously tested, with no change in the nature of the model.

Ashby performed a number of experiments on the Homeostat, both to prove that it could do what he had hypothesized it could, and to discover what else it was capable of. In terms of demonstrating that it was a valid model of the nervous system, Ashby replicated a number of experiments on his Homeostat that were previously performed on living systems. Among these was an experiment by Sperry in which the muscles in the arm of a monkey had been surgically cut, swapped around and reconnected such that they would now move the arm in the opposite direction than they previously had. In the experiment, the monkey relearns quite quickly how to control its arm in this new configuration. To replicate the experiment, Ashby used a switch on the front of a Homeostat unit which reversed the current passing through it. He then noted in the trace of the needle's movement, the exact opposite reaction to an incoming current as before, resulting in instability. However, after a short search by the Uniselector, the unit was able to stabilize once again (Ashby, 1952: 105-7). This sort of experiment was of the most basic kind.

More complicated experiments were also performed. These included conditioning experiments in which the needle was manually manipulated by the experimenter as a form of "punishment" (Ashby, 1952: 114). Ashby also describes an experiment in which he changes the rule governing the punishment systematically, sometimes enforcing one rule, and sometimes the other. From this he observes that the device is able to adapt to two different environments simultaneously (Ashby, 1952: 115). He also made new discoveries of the machine's capabilities through these explorations. These included ways of materially manipulating the machine that were made possible by its design, but were not intended features of the design. Most significant among these was the possibility of tying together with string, or binding with a stiff rod, the needles of two of the units. The result was to force them to find an equilibrium under the strange constraint of their material entanglement, as well as their electronic couplings (Ashby, 1952: 117). The Homeosat was quite capable of finding such an equilibrium. The other aspect of the material device that was not foreseeable on paper was its particular temporal dimensions–how quickly the Uniselector flipped, how quickly it found a new equilibrium, etc. Ashby also experimented with the delay in timing between the unit's behavior and the administration of punishments in conditioned training (Ashby, 1952: 120). These aspects of the machine were certainly incidental to its conception, but quite significant to real operations and provided the basis for further exploration of the Homeostat and its behavioral repertoires.

Ashby went on to develop a much more sophisticated synthetic brain, the Dynamic and Multistable System, or DAMS. While he had great hopes for DAMS, the technical design, cost, and above all the complexity of behavior that the device exhibited would frustrate Ashby for years. As Pickering (forthcoming) has recounted in his analysis of the ill-fated machine, there was also a clearly developmental aspect of his work on DAMS. Beginning with his great expectations for the device, Ashby's notebooks reveal the various causes of his frustration: from the problem of having to devise a clearer notion of "essential variables" than had sufficed for the Homeostat, to the technical difficulties of building such a large and complicated device, to the costs and his lack of research support funds, to Ashby's inability to comprehend the patterns of behavior exhibited by the complex device. Through these emerging developments, his understanding of the brain and what he was trying to do with DAMS evolved dynamically over time. It is clear that he was not certain how the device would behave specifically, but only sought a general sort of performance, yet what he got from it was still frustratingly complex, so much so that after nearly seven years of work, he abandoned it, and largely edited it out of the literature.

Walter's Tortoises were also subject to a number of experiments and explorations, and yielded unexpected results. Primarily the unexpected results were due to interactions between the robots based on their lights and light sensors:

> Some of these patterns of performance were calculable, though only as types of behaviour, in advance, some were quite unforeseen. The faculties of self-recognition and mutual recognition were obtained accidentally, since the pilot-light was inserted originally simply to indicate when the steering-servo was in operation. . . . The important feature of the effect is the establishment of a feedback loop in which the environment is a component. (Walter, 1953: 117)

What seems most clear from Walter's various discussions of the Tortoises is that he believed they were physiological models of biological phenomena. Of course, he was not claiming that the mechanisms which produced the behaviors in animals were identical to, or even isomorphic to, the circuits of the Tortoises, but just that those behaviors could be produced by simple circuits which employed feedback and modulation mechanisms similar to neural circuits – *i.e.* that they shared some relevant functional properties. An important scientific insight lay in the fact that very simple circuits could produce such complex and interesting phenomena provided that they instantiated these functions:

> The electronic "tortoise" may appropriately be considered as illustrating the use of models. But models of what? In the first place they must be thought of as illustrating the simplicity of construction of the cerebral mechanisms, rather than any simplicity in their organization... In short: reality may indeed be more complex than the model, but it is legitimate to consider that it may be equally simple. In a matter of which we know so little such a hypothesis is not without importance. (de Latil, 1957: 227-8).

The Tortoises were thus used as a basis to argue against those who hypothesized that neural circuits were necessarily or interminably "complex" because the behavior of organisms was complex.

Walter stopped short, however, of saying that these qualities were purely in the eye of the beholder. Instead, the phenomena of life and mind were posited to subsist in the negative feedback loops which held the organism in various stable configurations even as it moved through a dynamic environment. For these biologically-inspired and inspiring machines to become legitimate scientific models, they would have to do more than appear lifelike and evoke curiosity and wonder.

Despite the fact that they lacked a detailed theory or explained any specific set of data, these synthetic brain models produced useful and illuminating contributions to brain science. What becomes clear from looking at the history of these electronic models of the brain in the 1940s are some of the ways in which they served as mediators between relatively vague theories, and informal and often impressionistic observational data, and ultimately served as a stepping stone to more formal theories and rigorous observations. In short, they did this by providing models around which new theorization and observation could be organized. As a result, these rather simplistic models served as mediators between initially vague and unrelated theories, and more concrete and coherent theories of brain organization and behavior which followed.

The Tortoises and Homeostat succeeded in doing this because of their ability to demonstrate and communicate scientific knowledge in practical and accessible ways. These are pragmatic elements of epistemology, which is not simply a matter of truth and justification, but also depends upon pragmatic elements of the retention, transmission, and utilization of knowledge. This is what makes models so crucial to the epistemic efficacy of scientific knowledge. Such issues do not arise in traditional philosophy of science, which considers only the individual mind seeking justified theories, but do become crucial once we take a view of knowledge production as a social practice, in which multiple agents must generate knowledge together and transmit it to others. In these processes, issues of communication are not isolated from issues of knowledge, but are instead crucial to it. It is in this regard that the ability of a model to store and transmit knowledge, as image, idea and practices, become relevant. Ashby and Walter were both self-consciously aware of this fact.

It should also be clear that these examples of working models employ the synthetic method, and succeed scientifically in virtue of being mediators of a particular sort. Both the Homeostat and Tortoises are examples *par excellence* of the synthetic method–machines built according to specific principles of construction to offer a proof by example that a mechanistic approach to the brain could describe mechanisms that produce plausible behaviors. In the case of the Homeostat, it was shown that merely seeking equilibrium points and stability through random trial-and-error searches could lead to adaptive behavior. Whether a "Uniselector" was exactly the same mechanism that real creatures used to adapt to their environments was not at issue because of the level of abstraction at which the model operated. But it succeeded in forging a powerful conceptual linkage between learning and random search which still drives much of the research in AI and machine learning. In this sense the model was successful *because* it was a mediator between theories in biology (*e.g.* homeostasis) and theories in engineering (*e.g.* random search).

The Tortoises, with their elaborated set of behaviors demonstrate another aspect of how working models act as mediators within the synthetic method. Once a few major proofs by demonstration are achieved, it becomes necessary to explore and extend the working model. Walter did this both by extending the model to new psychological phenomena by observing these phenomena in the behavior of the Tortoises, and also by improving and extending the machinery of the working model. He built versions of the Tortoise which incorporated adaptive learning mechanisms, called CORA and IRMA (Walter, 1953). Using these mechanisms he "trained" his robots to respond in certain ways to a whistle by building an association between the whistle and another stimulus, much like Pavlov's dogs. Here the synthetic method works by drawing on the constraints of the real world – we could imagine all sorts of mechanisms that "might" produce such-and-such behavior, but here is one that actually *does* in this working model. In such cases the model is mediating in the more traditional sense, between theories and empirical data. But the empirical data are the behavior of the model itself, and the extension of the range and similarity of those behaviors to the target phenomena, animal behaviors, further reinforces the mediating bridge between theories, in this case psychological theories of behavior and interacting feedback control mechanisms.

## Conclusions

> What we have just reduced to absurdity is any prospect of reproducing all its elaboration of units in a working model. If the secret of the brain's elaborate performance lies there, in the number of its units, that would be indeed the only road, and that road would be closed. But since our inquiry is above all things a question of performance, it seemed reasonable to try an approach in which the first consideration would be the principles and character of the whole apparatus in operation. (Walter, 1953: 107)

As simulations, or stand-ins, working models were not completely new to science, but they were new to the brain sciences. The brain and mind sciences faced a challenge in the first half of the 20th century to become "empirical" through the use of the experimental method and to distance themselves from "metaphysical" forms of explanation (Gardner, 1987). Physics was at the time considered to be the science which other sciences should emulate methodologically, especially if they wanted to bolster their empirical legitimacy. There were two severe difficulties in doing this which were ultimately overcome through the synthetic method. The first was a requirement for the observability of the data-generating phenomena upon which theories were based. The second difficulty was that real functioning brains were very difficult to work on for technical, methodological and ethical reasons. The result was a behaviorist psychology that completely ignored the inner workings of the brain, and brain sciences that focused on the physiology of single or small numbers of cells, or the gross anatomy of mostly dead brains. The exception to this was psychiatry, which confronted sick and damaged brains on a regular basis, and managed to intervene in rather drastic ways upon them, in the hope of both curing the unhealthy brain and understanding the healthy brain.

As empirical experiments, working models offered a new basis for proceeding in areas of brain science that were otherwise difficult to address using the analytic "controlled experiments" method of physics because individual variables could not be so easily isolated. Despite efforts to shore up the epistemic basis of the brain sciences, the resulting approaches did not quite manage to integrate the various levels of analysis in a compelling way. It was the cybernetic brain models which managed to do this. But why should we think that building working models should be more useful than experimenting on neurons directly, or with the conditioned behavioral patterns of living animals?

This might be read as the "paradigm shift" from behaviorism and traditional neurophysiology to a computational cognitive neuroscience. It might also be possible to view this history as instead seeking a rigorous mathematical and mechanical account, and as a consequence of this arriving at theories which applied equally to the organic and inorganic. But there is another aspect of the use of these models that is a central theme of this article. It is the notion that these models acted as bridges or mediators between existing theories at different levels of analysis. Cordeschi acknowledges that working models established a new intermediary level of analysis between behavior and physiology. However, he emphasizes the autonomy of this new level of analysis, rather than its constructive and mediating aspects:

> The earliest behavioral models, designed as simple physical analogs, began to suggest that there might exist a *new* level for testing psychological and neurological hypotheses, one that might coexist alongside the investigations into overt behavior and the nervous system. This was the core idea of what Hull and Craik had already called the "synthetic method," the method of model building. This idea comes fully into focus with the advent of cybernetics, and especially when the pioneers of AI, who wanted to turn their discipline into a new science of the mind, found themselves coming to grips with the traditional sciences of the mind–psychology and neurology–and with their conflicting relationships. That event radically affected the customary taxonomies of the sciences of the mind. (Cordeschi, 2002: xviii)

The synthetic method was thus one way of establishing working models as mediators, mediators between theories espoused by very different disciplinary traditions, and mediators between very different sorts of phenomena and material performances.

The issue facing the brain sciences was how to proceed in devising and testing hypotheses about the biological basis of psychology. The first step in doing this was to give a plausible account that bridged low-level physiological mechanisms and high-level behavioral mechanisms. But there is a way of viewing the synthetic brains in which they might again appear puzzling. One can arrive at this puzzlement by noting that even while the brain sciences were seeking to construct a sound empirical basis for their work, they did this by positing a new level of analysis of the mind that was not itself directly observable. From this perspective, the synthetic brains are models of *mental functions*, not behavior or neurons, but rather functions that are not themselves directly observable. In fact, it is hard to conceive of just what entities these models were supposed to be modeling. The way out of this puzzle is to understand the empirical power of the mediation that they were performing between different levels of analysis. The two levels of analysis between which these new brain models tried to span were the traditional sciences of behavior and neurons, and in doing this they were able to combine the empirical force of both fields to bolster the hypothesized bridge between them. Without an existing theory of the interaction of the principle levels of analysis in the brain sciences, it was the particular advantage of the working synthetic brain models that they instantiated certain theories about *neurons* and were able to display characteristic *behaviors* as a consequence. It was the models themselves which mediated between the levels, not theoretically, but through their working and exhibiting behavioral phenomena that could stand in similarity relations to animal behavior.

Even if we take a view of science as ultimately about "knowing that," rather than "knowing how," working models are a practical means to knowledge and may very well be dispensable once the fruits of knowledge are collected. Indeed, it can often be difficult to recognize and express just how models have exerted their agency from this perspective. Yet, the moment we stop to ask why certain techniques and technologies are in place, and how they arrived at the forms they take, we would be at a loss to explain these in the absence of the history of material culture. In principle, no technological solution in use is the only one possible, and only rarely is it an optimal solution. Sometimes the technology in use is the best available, where the criteria of choice and the alternatives to choose among have evolved over time and in interaction with one another. And so history is the best way to get at such questions, and there is a sense in which every technological artifact is an archive of its own history–though it may have been subjected to much cleansing and scrubbing to remove its intrinsic historical traces. At the very least, we better understand a technology by knowing its history, as well as its structure.

It is clear that synthesizing and experimenting on working models can achieve many kinds of mediation useful to scientific progress. Working models have received far less attention in the philosophy of science literature than theoretical models have. Yet it seems clear that the pragmatic virtues of models function in many areas of scientific practice. It remains for further study to see to what extent working models can be found in other areas of science, and what roles they play as mediators in those areas. Taken broadly enough, one can find working models all over science. Yet it seems that the working models in each science are more or less unique to the science in which they are found. Indeed, it is their specificity to the scientific practices in which they are embedded that makes them both unique and interesting as a means to studying those practices. In this sense it may be undesirable to seek out any "general theory" of working models. Rather it seems more opportune to view these models as an obvious point of entry into studying the material culture of science. By asking such questions as "Why are the models built? How are they built? What kinds of experiments are performed? Which models are kept and which are discarded? How are they copied and developed over time and propagated through space?" we might hope to get at the very essence of the material culture of scientific models.

## References

Ashby, W. R.
1952 Design For a Brain. New York, NY: John Wiley and Sons.
1962 "Simulation of a Brain." Pp. 452-466 in Borko (ed.), Computer Applications in the Behavioral Sciences. New York, NY: Plenum Press.
1972 "Analysis of the System to be Modeled." Pp. 94-114 in Stogdill (ed.), The Process of Model-Building in the Behavioral Sciences. New York, NY: W. W. Norton.

Aspray, W.
1990 John von Neumann and the Origins of Modern Computing. Cambridge, MA: MIT Press.
Cartwright, N.
1983 How the Laws of Physics Lie. Oxford, UK: Clarendon Press.
Cordeschi, R.
2002 The Discovery of the Artificial: Behavior, Mind and Machines Before and Beyond Cybernetics. Dordrecht, The Netherlands: Kluwer.
de Latil, P.
1957 Thinking by Machine: A Study of Cybernetics. Y. M. Golla (trans. from French edition of 1956) Boston, MA: Houghton Mifflin Company.
Gardner, H.
1987 The Mind's New Science: A History of the Cognitive Revolution. New York, NY: Basic Books.
Giere, R.
1999 Science Without Laws. Chicago, IL: University of Chicago Press.
Hacking, I.
1983 Representing and Intervening: Introductory Topics in the Philosophy of Natural Science. Cambridge, UK: Cambridge University Press.
1992 "Do Thought Experiments Have a Life of Their Own?", in A. Fine and M. Forbes and K. Okruhlik (eds.) PSA 1992, 2: 302-310.
Hayward, R.
2001 "The Tortoise and the Love-Machine: Grey Walter and the Politics of Electroencephalography." Science in Context 14 (4): 615-641.
Hesse, M.
1966 Models and Analogies in Science. Notre Dame, IN: University of Notre Dame Press.
Kuhn, T.
1962 The Structure of Scientific Revolutions. Chicago, IL: University of Chicago Press.
Minsky, M.
1961 "Steps Toward Artificial Intelligence." Proceedings of the Institute of Radio Engineers, 49: 8-30.
Morgan, M. & Morrison, M. (eds.)
1999 Models as Mediators: Perspectives on Natural and Social Science. Cambridge, UK: Cambridge University Press.
Pickering, A.
1995 The Mangle of Practice: Time, Agency, and Science. Chicago, IL: The University of Chicago Press.
Forthcoming (Unpublished manuscript, February 2006.) Ontological Theatre: Cybernetics in Britain, 1940-2000. Chapter 4: Ross Ashby; Psychiatry, Synthetic Brains and Cybernetics. pp. 1-80.
Schaffer, S.
1994 "Machine Philosophy: Demonstration Devices in Gregorian Mechanics." Osiris 2nd series Instruments 9: 157-182.
Stafford, B.
1991 Body Criticism: Imaging the Unseen in Enlightenment Art and Medicine. Cambridge, MA: MIT Press.
Teller, P.
2001 "Twilight of the Perfect Model Model." Erkenntnis, 55(3): 393-415.
van Fraassen, B
1980 The Scientific Image. Oxford, UK: Clarendon Press.
Walter, W. G.
1953 The Living Brain. Harmondsworth, Middlesex, UK: Penguin Books.
Winsberg, E.
2003 "Simulated Experiments: Methodology for a Virtual World." Philosophy of Science 70 (December 2003): 105-125.
Yaneva, A.
2005 "Scaling Up and Down: Extraction Trials in Architectural Design." Social Studies of Science, 35(6): 867-894.

Peter Asaro
Gallery of Research
Austrian Academy of Sciences
Vienna, Austria
peterasaro@sbcglobal.net

Brain activity transitions between healthy states, including stages of sleep, restful and aroused waking, as well as pathological states such as epilepsy, coma, and unresponsive wakefulness syndrome. From such a diversity of brain states, phenomenological categories encompassing similar spatio-temporal activity patterns can roughly, but usefully, be defined: unconscious (e.g., sleep and anesthesia) and conscious (e.g., waking and dreaming) brain states.Â  The average Vm of the network, the Vm of a randomly chosen neuron, and the raster plot of the network are shown.Â  In this paper, we briefly reviewed work on the measurement and modeling of brain states at different scales, from single neurons to cell assemblies and global brain activity, considering both spontaneous and evoked dynamics.