

Understanding Endogenously Active Mechanisms: A Scientific and Philosophical Challenge

William Bechtel, Department of Philosophy, Center for Chronobiology, and
Science Studies Program, University of California, San Diego

Abstract

Although noting the importance of organization in mechanisms, the new mechanistic philosophers of science have followed most biologists in focusing primarily on only the simplest mode of organization in which operations are envisaged as occurring sequentially. Increasingly, though, biologists are recognizing that the mechanisms they confront are non-sequential and the operations nonlinear. To understand how such mechanisms function through time, they are turning to computational models and tools of dynamical systems theory. Recent research on circadian rhythms addressing both intracellular mechanisms and the intercellular networks in which these mechanisms are synchronized illuminates this point. This and other recent research in biology shows that the new mechanistic philosophers of science must expand their account of mechanistic explanation to incorporate computational modeling, yielding dynamical mechanistic explanations. Developing such explanations, however, is a challenge for both the scientists and the philosophers as there are serious tensions between mechanistic and dynamical approaches to science, and there are important opportunities for philosophers of science to contribute to surmounting these tensions.

The new mechanistic philosophers of science have abandoned the attempt to shoehorn biological explanations into nomological models and have focused instead on articulating the practices of biologists who appeal to mechanisms to explain biological phenomena. Most of these accounts have followed the practice, characteristic of biology through much of the 20th century, of treating mechanisms as sequentially organized reactive systems (Machamer, Darden, & Craver, 2000, for example, characterize mechanisms as “productive of regular changes from start or set-up to finish or termination conditions”). The result is what I characterize as a *basic account of mechanistic explanation*. But philosophers of science must remain attuned to ongoing developments in these sciences. In a host of fields biologists are recognizing that the mechanisms they confront are neither sequential nor reactive but employ non-sequential organization of non-linear interactions in open systems to generate endogenous activity. Understanding how biological organisms are endogenously active and how that affects their responses to stimulation is a challenge not just for biologists but also for philosophers of science, especially the new mechanists, as it requires integration of new tools, those of computational modeling, with the more traditional strategies for decomposing mechanisms. In this paper I present examples of recent biological research addressing endogenously active mechanisms and explore how relating the tools of mechanistic science with approaches for modeling dynamical systems provides a framework of *dynamic mechanistic explanations* that addresses the scientific and philosophical challenge.

1. The New Mechanistic Philosophy of Science Needs Dynamics

While the idea of explaining a phenomenon by identifying the responsible mechanism was championed by Descartes, and widely invoked in biology as well as many other sciences in the centuries since, it was eclipsed in 20th century philosophy of science by an account that viewed explanation as involving deduction of a linguistic characterization of the phenomenon to be explained from laws of nature and initial conditions (Hempel, 1965). Biologists, however, only infrequently appeal to laws and those are typically laws of physics or chemistry (Weber, 2005). Moreover, the primary effort in developing explanations in biology is not identifying the laws, but discovering and describing the constitution of the specific mechanism responsible for the phenomenon. This involves identifying its parts and the operations they perform and determining how these are organized to produce the phenomenon of interest. In the past couple decades several philosophers focused on biology have developed accounts of mechanistic explanation that attempt to reflect these explanatory endeavors of biologists (Bechtel & Richardson, 1993/2010; Glennan, 1996, 2002; Machamer et al., 2000; Bechtel & Abrahamsen, 2005; Darden, 2006; Craver, 2007). Their efforts have been characterized as constituting a new mechanistic philosophy of science.¹

Central to mechanistic explanation as it has been pursued in biology is the assumption that the behavior of mechanisms is to be understood in terms of the operations performed by their parts and that therefore it is essential to decompose mechanisms into their parts and operations. The ability of parts to perform operations is determined by their internal properties; whether they perform these operations is determined by the inputs they receive. Mechanistically oriented scientists have identified an extremely powerful set of tools for decomposing mechanisms, for example, by inhibiting or stimulating proposed operations in a mechanism and analyzing their effects on the overall mechanism to determine their specific contribution. Identifying how these parts are organized, however, has proven a good deal more challenging. Mechanistic research often begins by associating a whole phenomenon with a single part of the responsible system (e.g., Broca's identification of the capacity for articulate speech with a region of left prefrontal cortex or Buchner's identification of fermentation with an enzyme he designated *zymase*). In Bechtel and Richardson (1993/2010) we characterized this as the strategy of *simple* or *direct localization*, but noted that in practice it typically fosters research that leads to its own supplanting as researchers either identify other components that play fundamental roles in generating the phenomenon or discover that the single component actually contains multiple parts that carry out different operations (e.g., enzymes that catalyze a variety of reactions that all figure in fermentation). Although Richardson and I characterized the step of identifying multiple components of a mechanism as *complex localization*, in fact research usually begins by positing the simplest arrangement in which multiple parts are organized to generate the phenomenon—a sequential arrangement in which the product of one operation is provided as an input to the next operation, which transforms it and passes it to yet another operation, as in an assembly line. This is the sort of organization that is assumed in Machamer, Darden, and Craver's characterization of a mechanism as “productive of regular changes from start or set-up to finish or termination

¹ Important predecessors of the new mechanistic philosophy of science are Wimsatt (1976, p. 671), who argued that “at least in biology, most scientists see their work as explaining types of phenomena by discovering mechanisms . . .” and Salmon (1984), who defended a “causal-mechanical view” that drew out the importance of causal relations but said little about what renders a set of causal processes into a mechanism.

conditions.” When mechanisms are organized in this way, basic mechanistic explanations, in which one characterizes the functioning of the overall mechanism qualitatively in terms of the contribution of each of its parts and mentally rehearses (simulates) how it produces the phenomenon, suffice.

Historically evolving systems, however, are not limited to such sequential arrangements of parts; connections are spontaneously or opportunistically added between components, resulting in integrated networks of components. Often this involves connections through which operations later in what might be envisaged as a sequence send products (signals) back to operations envisaged as earlier in the sequence. When individual operations are affected by what is happening in multiple other parts of the system, some of which they affect through their own operation, and these processes occur on different timescales, the ability of qualitative mental simulation to determine the behavior of the overall system is severely compromised. This is especially true when, in addition to non-sequential organization, the operations within the mechanism are nonlinear when described mathematically and there is a flux of free energy through the mechanism. In such circumstances, mechanisms can exhibit complex behavior, including oscillations and even chaos.

Scientists in a variety of fields of biology are increasingly coming to recognize that mechanisms organized in the manner just described are endogenously active—they generate activity even when they are not being supplied with what are generally taken to be their inputs or start-up conditions. Among the biological phenomena that can result from endogenous activity in the responsible mechanisms are action potentials in neurons, contractions of heart muscles, locomotion of animals, and mind wandering (for examples and discussion, see Abrahamsen & Bechtel, in press). Endogenous activity, however, is often rendered invisible by the investigatory strategies researchers pursue. In many fields in which there is variability in the recorded behavior, researchers focus on the mean change in response to a stimulus and employ the variability in the response primarily in tests for the statistical significance of differences between means produced under different circumstances (e.g., the absence of a stimulus or presence of an alternative one). While variability in the signal recorded is noted, it is generally treated as noise that renders it difficult to extract what is regarded as the signal that reflects the response to the stimulus. In fact, such noise often reflects the endogenous activity of the system. Far from being in a constant state, the mechanism varies over time and this has consequences for the activity that might be evoked by what are usually taken as the inputs to the mechanism. To identify and study this endogenous activity requires a different approach than that which reveals the parts and operations of a mechanism—researchers must examine the behavior of the mechanism across time, employing time-series analysis, and then analyze the patterns of change using tools of mathematical modeling and dynamical systems theory.

Explaining the behavior of endogenously active mechanisms requires what Abrahamsen and I (Bechtel & Abrahamsen, 2010) refer to as *dynamic mechanistic explanations*, explanations that take into account the parts and operations of the mechanism, their spatial organization, and the “patterns of change over time in properties of its parts and operations” that result in the orchestrated behavior of the mechanism. Mathematical models that employ differential equations to describe the individual operations are employed to simulate the behavior of the mechanism through time. Determining appropriate parameter values is often a major part of the modeling enterprise. In many cases, the models show that some parameter values result in an initial

transient before the mechanism settles into a stable state. But with other parameter values, more complex patterns of change are generated in the models and researchers must then investigate whether these parameter values correctly describe the actual mechanism.

Endogenously active mechanisms often exhibit the simplest of these more complex patterns, oscillatory behavior. In the following section, I discuss the types of mechanisms that can explain endogenous oscillatory phenomena that are ubiquitous in biology. In a subsequent section I consider strategies for understanding complex phase relations that can be established between oscillatory mechanisms before returning in the final section to the challenges both developments present for both scientists themselves and philosophers of science.

Biological Oscillations and the Responsible Mechanisms

Some oscillatory processes in biology are widely recognized—heart rhythms and ovulation in mammals, fluorescent flashing of fireflies, annual cycles of flowering in plants and hibernation in some animals, etc. But in fact oscillatory processes are ubiquitous in biological organisms. Cells exhibit cycles of division and growth. Oxidative metabolism and sleep exhibit cycles of approximately 90 minutes. Brain processes exhibit oscillations at a number of frequencies, from relatively slow oscillations with periods greater than 10 seconds, detected with fMRI, to much faster oscillations (1-80 Herz), detected by EEG or implanted electrodes. Conventionally, biological oscillations are differentiated into ultradian (those with a period well less than 24 hours), circadian (those with an approximately 24 hour period), and infradian (those with a period of greater than 24 hours).

In this paper I will focus on circadian rhythms as extensive research has resulted in quite rich understanding of them. These oscillations are entrainable to the daily cycle on the planet by cues such as light and temperature. But they are maintained in the absence of such cues, a condition known as *free-running*, during which they exhibit periods deviating only slightly from 24 hours (hence, the name *circa* [about] + *dies* [day]). They are found in all orders of life, from bacteria to fungi, plants, and animals. And they regulate a vast range of physiological and behavioral phenomena, including, in humans, sleep and attention, physiological strength and mental reaction times, heart rate and basic metabolism, and hunger, thirst, and waste excretion.

Oscillatory mechanisms such as pendulums are well known in physics, where the opposition of driving and resisting forces with time delays can generate oscillations. Often oscillations arise where they are not desired, as when negative feedback is employed with the objective of maintaining constant values for regulated variables but ends up producing oscillations instead. For example, a common household thermostat is designed to keep a room at a specified temperature, but results in an oscillation between a slightly lower and a slightly higher temperature. Considerable engineering work has gone into developing procedures to minimize the oscillatory range. But in biology such oscillations are sometimes useful, and research has been devoted to how oscillations may be maintained. For example, when the lac operon was identified as a feedback mechanism at the genetic level in bacteria (wherein a gene product figures in regulating gene expression), Goodwin (1965) undertook an investigation of when such mechanisms might produce sustained oscillations. He determined that in addition to feedback and a continual supply of energy, at least one of the reactions had to be non-linear when described mathematically. Goodwin's account of the conditions under which oscillation occurred

influenced modeling of mechanisms for creating circadian rhythms once empirical research began to provide clues to the basic mechanism.

Konopka and Benzer (1971) provided the first clue to the underlying mechanism when they identified a *Drosophila* gene, which they named *period* (*per*), whose various mutants exhibited shortened or lengthened periods or became arrhythmic. In the 1980s it became possible to measure the changing concentrations of *per* mRNA and the resulting protein (PER) and both were found to oscillate, with *per* mRNA reaching peak concentrations several hours in advance of the protein. This led Hardin, Hall, and Rosbash (1990) to propose a feedback mechanism in which *per* was transcribed into mRNA in the nucleus, the mRNA was transported to the cytoplasm to synthesize the protein PER, and the protein was in turn transported back into the nucleus where, in some at the time unspecified manner, it inhibited further transcription (see Figure 1).

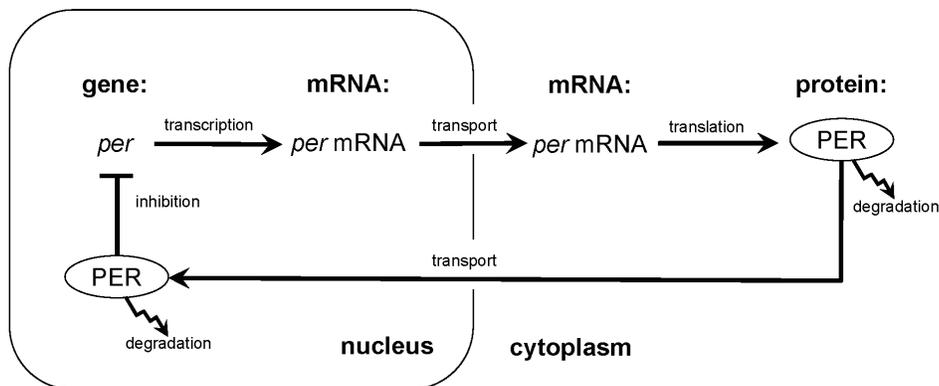


Figure 1. Hardin, Hall, and Rosbash's (1990) proposed feedback mechanism for generating circadian oscillations in *Drosophila*.

Since molecules of PER degrade over time it seemed plausible that the proposed mechanism would generate oscillations, but it was possible that it would instead settle into an equilibrium state. While further mechanistic research (decomposing the system into components and discovering how they were organized in the cell) was needed to resolve questions such as how PER could inhibit *per* transcription, it could not reveal the temporal dynamics of the mechanism. Pursuing that goal required a different approach, one that focused on how the proposed mechanism would behave in real time. That is, the mechanistic explanation provided only the starting point for a dynamic mechanistic explanation. Goldbeter (1995) took the next step by offering a mathematical model of the behavior of Hardin et al.'s circadian mechanism, comprising five differential equations. Figure 2 shows the first of these equations and illustrates how its terms and variables correspond to properties of certain parts and operations in the mechanistic account.

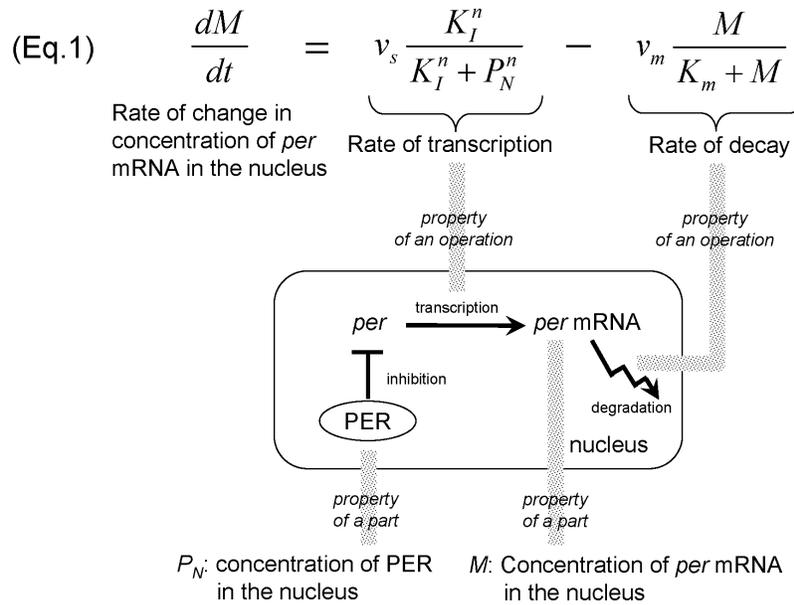


Figure 2. A dynamic mechanistic explanation. Equation (1) in Goldbeter’s (1995) model shown in relation to the relevant portion of Hardin et al.’s proposed circadian mechanism (Figure 1). A property of certain parts of that mechanism (the concentration of *per* mRNA and of its protein, PER) directly correspond to the variables M and P_N . The salient property of two operations—their rate—corresponds to the two terms of the equation. These terms include several parameters: v_s represents the maximum rate for the accumulation of M , K_I is a threshold constant for inhibition, n is the Hill coefficient indicating the minimum number of cooperating molecules required to achieve inhibition, v_m is the maximum rate for the degradation of M , and K_m is the Michaelis constant for the degradation reaction.

Using biologically plausible parameter values in the mathematical model of its dynamics, Goldbeter succeeded in demonstrating that such a mechanism could indeed maintain oscillation. In fact, the system of equations generated what is known as a limit cycle. By plotting the successive states of the system in phase space (with mRNA and protein concentrations as the two axes), the system is seen to approach the cycle (dark oval in Figure 3) as a limit and, if ever perturbed from this cycle, to return to it. The oval represents the continued oscillations such a system would exhibit were the limit ever reached.

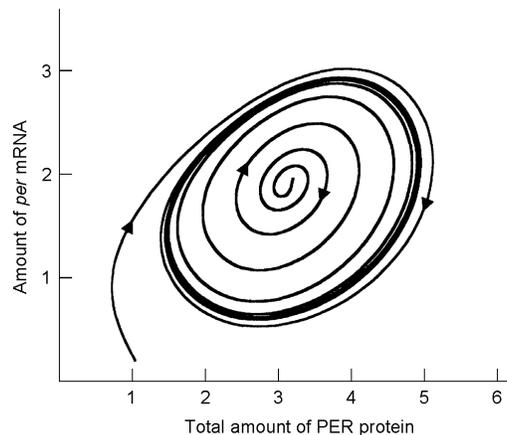


Figure 3. Limit cycle generated by Goldbeter’s (1995) mathematical model.

Even at the time this mechanism was proposed and being modeled, researchers suspected there were more components (one indicator was that PER lacks a DNA binding region and so something else must mediate its inhibition of its own transcription). In short order researchers identified numerous other *Drosophila* clock genes comprising positive and negative feedback loops (several of these additional components are shown in Figure 4). Although one can try, especially by appealing to such a figure, to rehearse mentally the behavior of such a mechanism, it becomes very difficult to anticipate how the positive feedback loop generating CLOCK interacts with the negative feedback loop through which PER interacts with CLOCK, and how that interacts with CRY when light is present. Will the mechanism sustain oscillations, or settle into a stable state? Again, this was a question that required mathematical modeling to answer: a variety of models (Leloup & Goldbeter, 2000; Smolen, Hardin, Lo, Baxter, & Byrne, 2004) indicate that under biologically plausible parameters stable oscillations will result.

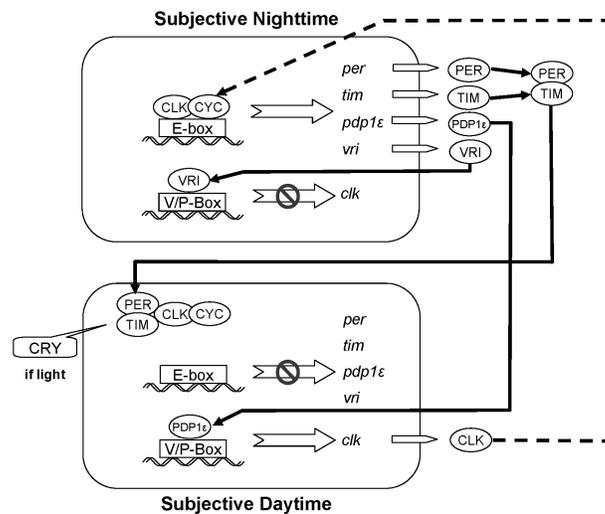


Figure 4. A more complete account of the *Drosophila* oscillator. The large open arrows indicate whether the promoter turns gene expression on or off. The smaller open arrows represent the combined processes of gene transcription in the nucleus, transport to the cytoplasm, and translation in the cytoplasm.

The primary function of these models is to determine how a mechanism involving interaction of multiple parts will behave. But they also can be used to address other questions. For example, is the mechanism robust to changes in parameter values, or are there parameter values that radically alter the behavior of the mechanism? Such questions are typically addressed by running the computational model under various parameter settings and analyzing the results (often by portraying them graphically). Smolen et al., for example, manipulated the parameters affecting CLOCK to determine whether the positive feedback loop was essential to generating sustained oscillations (in their models it did not appear to be so). Leloup and Goldbeter investigated parameters affecting the breakdown of PER and TIM, and found values that produced behavior corresponding to known sleep pathologies, thereby suggesting how the mechanism could explain these pathologies. The point to be emphasized is that such modeling provides understanding beyond that which is available from identifying the parts, operations, and organization of the

mechanism and mentally rehearsing its functioning. With mechanisms exhibiting complex behavior, modeling has a crucial role to play.

This brief exposition of the mechanism generating circadian rhythms in *Drosophila* has considered only research on the core mechanism for generating rhythms, and has not discussed either the processes by which these rhythms are entrained by light or those through which they serve to regulate physiological processes and behaviors of the fly. Although much has been learned in the past decade, there is still far less known about these processes than about the core mechanism. What is becoming apparent, though, both from research on flies and on mammals, is that these processes are themselves not sequential but involve complex feedback processes whereby the core mechanism regulates its own inputs, and what appear to be output systems affect the clock's own behavior (see Bechtel & Abrahamsen, 2009, pp., for details). This means that ultimately neither the core mechanism nor its input or output system can be fully understood independently of the others as the operation of components within the core mechanism both affect the operations performed elsewhere in the organism and are affected by those. This further undermines the prospect of understanding the mechanism in terms of a sequence of operations from start to termination condition. Fortunately, once they have developed a mathematical model for the core mechanism, modelers can add additional terms to the appropriate equations to characterize how operations occurring elsewhere affect, even in a nonlinear fashion, the components of the core mechanism itself. Thus, dynamic accounts of mechanism can succeed in understanding the oscillatory behavior produced in mechanisms for which simple sequential accounts fail.

Populations of Synchronized Oscillators in Complex Phase Relations

In this section I consider a further example in which synchronization of oscillators illustrates the need to develop dynamic mechanistic explanations. For this example I shift from *Drosophila*, in which a small number of ventral and lateral neurons constitute the core circadian mechanism, to mammals, in which a part of the hypothalamus, known as the suprachiasmatic nucleus (SCN), a structure consisting for 8,000-10,000 neurons on each side of the brain in mice, serves as the core mechanism.² Lesions to the SCN result in loss of circadian behaviors, and transplanting a donor SCN into the ventricles restores some of these circadian behaviors (Silver, LeSauter, Tresco, & Lehman, 1996). Initially it seemed plausible that individual SCN neurons all behaved the same, so that the behavior of the SCN was simply a collective effect of its individual neurons, but Welsh, Logothetis, Meister, and Reppert (1995) showed that when SCN neurons were dissociated in culture on a microelectrode array (which allowed for retention of “abundant functional synapses”), some neurons oscillated in antiphase with others and the periods of oscillation varied substantially between neurons (ranging from 21.25 to 26.25 hours, with a standard deviation of 1.25 hours). In contrast, the circadian behavior in organisms is much more regular, and indeed Welsh himself had previously shown that the overall circadian signal shows virtually no variability (Welsh, Engle, Richardson, & Dement, 1986). Subsequent research demonstrated that the variability was largely eliminated in non-dispersed explants in which

² There are some important differences between the basic circadian mechanism in *Drosophila* and mammals, but it is largely conserved. This conservation facilitated the identification of parts of the mammalian mechanism, and the discovery of new components in the mammalian mechanism also supported comparable discoveries in *Drosophila* (see Bechtel, 2009).

nearly complete connectivity is maintained (Herzog, Aton, Numano, Sakaki, & Tei, 2004). This reveals that regular time keeping emerges at the population level from coupling processes that synchronize highly variable component timekeepers.

Coupling of oscillators requires the transmission of a signal between oscillators, and research on possible coupling agents in the SCN identified vasoactive intestinal polypeptide (VIP) as the primary synchronizing agent, although GABA and gap-junctions may also be involved (Aton & Herzog, 2005). The process of coupling is complicated by the fact that there are at least two distinct regions of the SCN, a core and a shell (van den Pol, 1980). Differences between these regions include: (1) patterns of connectivity by which they receive inputs from other parts of the brain and send outputs to other brain areas and other organs, (2) only neurons in the core release VIP (those in the shell produce arginine vasopressin instead), (3) neurons in the shell oscillate in advance of those in the core, but (4) only those in the core maintain sustained endogenous oscillations.

The ability of oscillators to synchronize when a signal is transmitted between them has been known since the observations of Huygens, but over the 20th century it was also recognized that, depending on the character of the oscillators themselves and the timing of the signal between them, the resulting behavior can be extremely complex, sometimes involving toroidal oscillations, deterministic chaos, or coexistence of multiple attractors (Grebogi, Ott, & Yorke, 1987). Computational modeling is thus essential for understanding the effects of the release of VIP in synchronizing the oscillations of individual neurons in the SCN.

The first effort to model the coupling process, by Gonze, Bernard, Waltermann, Kramer, and Herzog (2005), did not take into account the anatomical details about the SCN, but simply explored whether generation of a compound such as VIP could synchronize independent mechanisms of the sort thought to be responsible for circadian rhythms. The modelers adapted a version of the Goodwin oscillator (mentioned above) to represent the generation and degradation of a single clock protein in each of 1000 modeled neurons and added equations describing the change in concentration of VIP as it was created by each oscillator and then degraded and calculating the mean of VIP concentration across the population. They then added a term to the equation describing the change in concentration of the clock protein that increased the rate of change proportional to the mean concentration of VIP. When the parameter in this term was set to 0, Gonze et al. obtained results much like those of Welsh et al. (the periods of different oscillators were highly variable), but when it was set to 0.5, the oscillators synchronized.

In this initial model, Gonze et al. assumed that VIP was produced by all SCN neurons and instantly and equally distributed to all others. This is one of the modes of network organization whose properties were investigated by mathematicians in the field of graph theory in the mid-20th century (for discussion and references, see Strogatz, 2001). Mathematicians also explored the properties of two other network structures, regular lattices and randomly connected networks were also explored, and deployed two useful measures for characterizing information flow, characteristic path length and the clustering coefficient. The characteristic path length is the mean of the shortest path between pairs of nodes and reflects how quickly information can be transmitted through the network. The clustering coefficient is the proportion of possible links in local neighborhoods that are actually realized and reflects how much specialized processing can be accomplished by cooperating nodes. Short characteristic path lengths and higher clustering

are desirable for information processing and are realized in fully connected networks. However, maintaining complete connectivity between all neurons in a network is metabolically very expensive. However, the other modes of organization initially explored each only exhibit one of the desirable features: regular lattices only allow high clustering whereas random networks only facilitate short characteristic path lengths.

The tradeoff between path length and clustering, however, is not inevitable. Watts and Strogatz (1998) investigated a new class of networks in which most connections are between local units but a few are long-distance. They found that these networks, which they termed *small-worlds*, exhibited both short characteristic path lengths and high clustering. They also demonstrated that a host of real-world networks, including networks of movie actors linked by co-appearances, the electrical power-grid of the Western U. S., and the neural network of the nematode worm *Caenorhabditis elegans*, exhibit small-world properties. In addition, they examined the functional properties of small-world networks, showing how they allow for rapid spread of infectious diseases, enable efficient problem solving in cellular automata, and reduce the likelihood of cooperation in iterated prisoner dilemma games. Of particular interest, they examined coupled phase oscillators and demonstrated that synchronization occurred almost as fast in small-world networks as in fully-connected networks. They speculated that the brain has a small-world architecture and that this could explain the synchronization of widely separated neurons in visual cortex. This speculation has been supported by connectivity matrices Sporns and Zwi (2004) developed based on published neuroanatomical data, including Felleman and van Essen's (1991) study of the macaque's visual cortex. Moreover, Sporns and Zwi showed that the brain networks exhibit characteristic properties of small-worlds: short characteristic path lengths and high clustering. (See Bullmore & Sporns, 2009, for additional analyses of structural and functional connectivity in brain networks.)

Might the SCN exhibit a small-world architecture? Vasalou, Herzog, and Henson (2009) investigated this question in a modeling study. They began with a previous effort of their own (To, Henson, Herzog, & Doyle, 2007) in which they adapted the model of mammalian circadian oscillations developed by Leloup and Goldbeter (2004) so that only some neurons maintained oscillations without synchronization and VIP diffused from each neuron in which it was generated. They modified the connections in the network to reflect small-world connectivity and demonstrated synchronization that approximated what was achieved in a totally connected network. They were also able to capture three other phenomena observed in experimental studies: with VIP (1) the percentage of oscillating neurons in the SCN rises from about 30% to nearly all, (2) the period is extended from approximately 22 to approximately 24 hours, and (3) the variability in periods is largely eliminated. Vasalou et al. emphasize the cost virtues of small-world networks over totally connected networks—they can achieve the same synchronization with fewer neural connections, which are very energetically expensive to develop and maintain. The researchers do not, however, pursue the possibility that the high clustering in small-worlds may generate other useful features, including the ability of localized clusters to maintain oscillations offset from others that may serve to regulate different functions that must be performed at different times of day (e.g., those linked to the onset of daylight and those linked to the end of daylight). Demonstrating that the SCN actually exhibits small-world organization and what virtues that offers requires further research.

What is noteworthy is that all of the research on how synchronization between oscillators might arise in the normally functioning SCN has been conducted in mathematical simulations. This reflects the fact that the knowledge sought about the dynamic behavior in a complexly organized mechanism goes beyond what basic mechanistic research alone could provide. Mechanistic research provided needed information about the component parts of the SCN and their individual behaviors and about the basic neural architecture found in the SCN, but was not itself able to determine the temporal behavior such a mechanism would produce. This is what mathematical modeling provides. The modeling efforts, though, do not eclipse mechanistic research. Rather, they built upon its results, and posed new questions about the realism of the models (e.g., about the plausibility of the parameters employed in the models and their robustness under various parameter changes) which themselves can only be answered by further mechanistic research.

Dynamics and the Future of Mechanistic Philosophy of Science

In the two previous sections I developed examples in which research on circadian rhythms has identified the non-sequential organization of non-linear operations that generate the rhythms and the processes by which oscillators with variable periods are synchronized through networks of connections to produce regular oscillations. I offer these not as eccentric examples but as exemplary of recent developments in biology. My point in presenting these examples is to show how biological understanding of these mechanisms requires supplementing the traditional basic approach of mechanistic explanation exemplified in biological practice and described in the philosophical accounts of the new mechanists. Traditional mechanistic modes of inquiry have been highly successful in identifying the components of the individual oscillatory mechanisms and the agents through which they couple. Such research, however, is not able to show how a mechanism with the proposed parts, performing the specified operations, and organized in a specified manner, will actually behave. This requires bringing to bear additional tools, those of mathematical modeling and dynamical systems analysis, to offer dynamic mechanistic explanations.

In arguing for the need for philosophy of science to attend to dynamic mechanistic explanation I am embracing the same naturalistic approach that initially gave rise to the new mechanistic philosophy of science. The new mechanism emerged as philosophers of biology noted the poor fit of traditional philosophical approaches to explanation to prominent instances of explanation in biology. But biology is a moving target. While mechanistic research directed at decomposing mechanisms into their parts and operations is still fundamental, biologists are increasingly recognizing and confronting the challenges of recomposing mechanisms and understanding the complex behavior that results when the organization is non-sequential. The recognition of the need to employ mathematical models and dynamical systems analysis to understand biological mechanisms is one of the factors contributing to the recent rise of systems biology (Noble, 2006; Boogerd, Bruggeman, Hofmeyr, & Westerhoff, 2007). Investigators are increasingly discovering that mechanisms that were once envisaged as operating sequentially involve feedback loops that support complex dynamical behavior. If philosophy of science is to characterize actual science, its models of explanation must accommodate the integration of dynamics into mechanistic explanations.

So far I have presented mathematical modeling as an approach that complements and extends basic mechanistic research. But integration of mechanistic and dynamical approaches is not easy

and can be challenging to bring about. Mechanistic research strategies require researchers to decompose mechanisms into their parts and characterize the properties of these parts independently of their role in the mechanism. A variety of experimental manipulations are employed to reveal the operations performed by the parts of the mechanism and it is assumed that the parts operate the same under these experimental manipulations as they do in the mechanism as found in nature. However, as a result of non-sequential organization, the operations of the parts are modulated by operations occurring elsewhere in the mechanism and these effects are typically altered in the experimental context. Far from being intrinsic to the parts, these operations are context sensitive. Such variability is expected in dynamical models, in which the parts are characterized in terms of variables—things that vary. So there is inherent tension in construing the parts both as having fixed properties, as envisaged in mechanistic research emphasizing decomposition, and being variable, as characterized in dynamical models. Yet, both approaches are needed in order to understand the endogenously active mechanisms found in biology—researchers need to identify parts in terms of their properties *and* recognize how these change over time, partly as a result of other changes in the mechanism and external factors impinging on the mechanism. In practice, biologists often proceed by first assuming that the parts and operations are stable across contexts and so describe them. These, however, serve as “first approximations” (Bechtel & Richardson, 1993/2010) which then must be revised in light of further research that reveals ways they are sensitive to context.³ Were the effects of context to swamp any attempt to track parts and operations, mechanistic explanation would fail (a point clearly articulated by Simon, 1969, in his discussion of near-decomposability). In much of biology, fortunately, it remains possible to identify parts and operations even as mathematical equations are employed to characterize how they are affected in various ways by being incorporated into complex systems.

The tension that arises between mechanistic decomposition and mathematical modeling is also reflected in philosophical positions. Some of the new mechanists (e.g., Craver, 2007) reject the proposal that mathematical models are explanatory. But in order to determine that a mechanism could explain a phenomenon, investigators must be able to show that it could generate the phenomenon. For relatively simple mechanisms, they can proceed sequentially, mentally representing each process of change, but I have stressed that this is not possible when the organization of the mechanism is non-sequential, the operations are non-linear, and there is a flux of free energy through the mechanism. Then the only tools for determining that the mechanism will generate particular behaviors are those of mathematical modeling and dynamical systems theory. Without them, mechanistic explanation is blind.

One of the factors that motivates Craver to deny that models, including computational models, are explanatory, is that models can be proposed that are not grounded in experimental research identifying the parts and operations incorporated in the model. He denies, for example, that Hodgkin and Huxley’s model of the action potential is explanatory since the key components of the mechanism, the ion channels, and the operations by which they opened and closed, had not

³ One might view the context-sensitive behavior of parts and operations as a reason to repudiate the project of decomposition. For reasons indicated below, I contend that the project of decomposition remains crucial to mechanistic biology. The context sensitivity of components can be accommodated by articulating the interaction between intrinsic features of components with processes external to them.

been discovered and the key parameters in the model resulted from fitting data on the voltage sensitive conductance of ions. They were not grounded in details of the mechanism. Likewise, many advocates of dynamical approaches are critical of mechanistic research, maintaining that developing mathematical characterizations of how variables change over time is sufficient for explanation (Chemero, 2000). Left ungrounded, the variables in dynamical equations cannot be related to the actual processes bringing about the phenomenon. The proffered explanation reflects only a possible way the phenomena might have been generated, and accordingly is empty.

Computational models, however, need not be so detached from the details of the mechanism. Modelers structured the equations in the models of circadian oscillations in *Drosophila* discussed above in light of what was known of the parts and operations of the responsible mechanism. Their exploration of parameters was less directly constrained, but they were still concerned to show that the parameter values that produced appropriate circadian behavior were biologically plausible. In modeling the network structure that would produce synchronization, researchers are constructing models that exceeded what is known of the connectivity of the SCN. But part of their objective is to provide guidance for further experimental investigations. The models are not floating free of the mechanism, and an important part of the evaluation of the adequacy of such models is whether they accurately describe the mechanism (see also Kaplan & Craver, in press). In these cases, dynamical accounts do not supersede mechanistic research, but provide understanding of how a mechanism with a given constitution actually behaves.

Conclusion

Both in science and in the philosophy of science, there is a challenge in integrating mechanistic research emphasizing decomposition and dynamical modeling in dynamic mechanistic explanations. But such integration is essential if science is to understand the endogenously active mechanisms that are found in living systems. Scientists in fields such as circadian rhythm research are increasingly discovering the need to integrate the two approaches and are attempting to do so, and philosophers can both learn from and potentially contribute to the ongoing efforts at integration in this and other domains of biology that confront endogenously active mechanisms. In doing so, philosophy of science can maintain the sort of productive engagement with science that has been one of its signature virtues in recent decades.

References

- Abrahamsen, A., & Bechtel, W. (in press). From reactive to endogenously active dynamical conceptions of the brain. In T. Reydon & K. Plaisance (Eds.), *Philosophy of behavioral biology*. New York: Springer.
- Aton, S. J., & Herzog, E. D. (2005). Come together, right...now: Synchronization of rhythms in a mammalian circadian clock. *Neuron*, 48(4), 531-534.
- Bechtel, W. (2009). Generalization and discovery by assuming conserved mechanisms: Cross species research on circadian oscillators. *Philosophy of Science*, 76, 762-773.
- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 421-441.

- Bechtel, W., & Abrahamsen, A. (2009). Decomposing, recomposing, and situating circadian mechanisms: Three tasks in developing mechanistic explanations. In H. Leitgeb & A. Hieke (Eds.), *Reduction and elimination in philosophy of mind and philosophy of neuroscience* (pp. 173-186). Frankfurt: Ontos Verlag.
- Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, 41(3), 321-333.
- Bechtel, W., & Richardson, R. C. (1993/2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Cambridge, MA: MIT Press. 1993 edition published by Princeton University Press.
- Boogerd, F., Bruggeman, F. J., Hofmeyr, J.-H., & Westerhoff, H. (Eds.) (2007). *Systems biology: Philosophical perspectives*. Amsterdam: Elsevier.
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3), 186-198.
- Chemero, A. (2000). Anti-representationalism and the dynamical stance. *Philosophy of Science*, 67(4), 625-647.
- Craver, C. F. (2007). *Explaining the brain: What a science of the mind-brain could be*. New York: Oxford University Press.
- Darden, L. (2006). *Reasoning in biological discoveries: Essays on mechanisms, interfield relations, and anomaly resolution*. Cambridge: Cambridge University Press.
- Felleman, D. J., & van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1-47.
- Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, 44, 50-71.
- Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, 69, S342-S353.
- Goldbeter, A. (1995). A model for circadian oscillations in the *Drosophila* Period protein (PER). *Proceedings of the Royal Society of London. B: Biological Sciences*, 261(1362), 319-324.
- Gonze, D., Bernard, S., Waltermann, C., Kramer, A., & Herzog, H. (2005). Spontaneous synchronization of coupled circadian oscillators. *Biophysical Journal*, 89(1), 120-129.
- Goodwin, B. C. (1965). Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation*, 3, 425-428.
- Grebogi, C., Ott, E., & Yorke, J. A. (1987). Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics. *Science*, 238(4827), 632-638.
- Hardin, P. E., Hall, J. C., & Rosbash, M. (1990). Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. *Nature*, 343(6258), 536-540.
- Hempel, C. G. (1965). Aspects of scientific explanation. In C. G. Hempel (Ed.), *Aspects of scientific explanation and other essays in the philosophy of science* (pp. 331-496). New York: Macmillan.
- Herzog, E. D., Aton, S. J., Numano, R., Sakaki, Y., & Tei, H. (2004). Temporal precision in the mammalian circadian system: A reliable clock from less reliable neurons. *Journal of Biological Rhythms*, 19(1), 35-46.
- Kaplan, D. M., & Craver, C. (in press). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective.
- Konopka, R. J., & Benzer, S. (1971). Clock mutants of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences (USA)*, 68, 2112-2116.
- Leloup, J.-C., & Goldbeter, A. (2000). Modeling the molecular regulatory mechanism of circadian rhythms in *Drosophila*. *BioEssays*, 22(1), 84-93.

- Leloup, J.-C., & Goldbeter, A. (2004). Modeling the mammalian circadian clock: Sensitivity analysis and multiplicity of oscillatory mechanisms. *Journal of Theoretical Biology*, 230(4), 541-562.
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67, 1-25.
- Noble, D. (2006). *The music of life: Biology beyond the genome*. Oxford: Oxford University Press.
- Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton, N.J.: Princeton University Press.
- Silver, R., LeSauter, J., Tresco, P. A., & Lehman, M. N. (1996). A diffusible coupling signal from the transplanted suprachiasmatic nucleus controlling circadian locomotor rhythms. *Nature*, 382(6594), 810-813.
- Simon, H. A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.
- Smolen, P., Hardin, P. E., Lo, B. S., Baxter, D. A., & Byrne, J. H. (2004). Simulation of *Drosophila* circadian oscillations, mutations, and light responses by a model with VRI, PDP-1, and CLK. *Biophys. J.*, 86(5), 2786-2802.
- Sporns, O., & Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics*, 2(2), 145-162.
- Strogatz, S. H. (2001). Exploring complex networks. *Nature*, 410(6825), 268-276.
- To, T.-L., Henson, M. A., Herzog, E. D., & Doyle, F. J., III. (2007). A molecular model for intercellular synchronization in the mammalian circadian clock. *Biophysical Journal*, 92(11), 3792-3803.
- van den Pol, A. N. (1980). The hypothalamic suprachiasmatic nucleus of rat: Intrinsic anatomy. *The Journal of Comparative Neurology*, 191(4), 661-702.
- Vasalou, C., Herzog, E. D., & Henson, M. A. (2009). Small-world network models of intercellular coupling predict enhanced synchronization in the suprachiasmatic nucleus. *Journal of Biological Rhythms*, 24(3), 243-254.
- Watts, D., & Strogatz, S. (1998). Collective dynamics of small worlds. *Nature*, 393(440-442).
- Weber, M. (2005). *Philosophy of experimental biology*. Cambridge: Cambridge University Press.
- Welsh, D. K., Engle, E. M. R. A., Richardson, G. S., & Dement, W. C. (1986). Precision of circadian wake and activity onset timing in the mouse. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 158(6), 827-834.
- Welsh, D. K., Logothetis, D. E., Meister, M., & Reppert, S. M. (1995). Individual neurons dissociated from rat suprachiasmatic nucleus express independently phased circadian firing rhythms. *Neuron*, 14(4), 697-706.
- Wimsatt, W. C. (1976). Reductive explanation: A functional account. In R. S. Cohen, C. A. Hooker, A. C. Michalos & J. van Evra (Eds.), *PSA-1974* (pp. 671-710). Dordrecht: Reidel.