Dynamic Mechanistic Explanation: Computational Modeling of Circadian Rhythms as an Exemplar for Cognitive Science

William Bechtel
Department of Philosophy and Interdisciplinary Programs
in Science Studies and Cognitive Science
University of California, San Diego

Adele Abrahamsen
Center for Research in Language
University of California, San Diego

Two widely accepted assumptions within cognitive science are that (1) the goal is to understand the mechanisms responsible for cognitive performances and (2) computational modeling is a major tool for understanding these mechanisms. The particular approaches to computational modeling adopted in cognitive science, moreover, have significantly affected the way in which cognitive mechanisms are understood. Unable to employ some of the more common methods for conducting research on mechanisms, cognitive scientists' guiding ideas about mechanism have developed in conjunction with their styles of modeling. In particular, mental operations often are conceptualized as comparable to the processes employed in classical symbolic AI or neural network models. These models, in turn, have been interpreted by some as themselves intelligent systems since they employ the same type of operations as does the mind. For this paper, what is significant about these approaches to modeling is that they are constructed specifically to account for behavior and are evaluated by how well they do so—not by independent evidence that they describe actual operations in mental mechanisms.

Cognitive modeling has both been fruitful and subject to certain limitations. A good way of exploring this is to contrast it with a different approach, one that involves more direct investigation into the internal parts and operations of the mechanism responsible for a phenomenon and tailors modeling to this mechanism. To do this we will focus on the phenomenon of circadian rhythms: the ability of the nervous system to regulate the activities of organisms, including human cognitive activities, on an approximately 24 hour cycle. Circadian effects on cognition generally have been ignored in cognitive science, but whether or not that is a desirable state of affairs is not relevant here. Rather, our goal is to use the increasingly prominent role of computational modeling in circadian rhythm research as a different type of exemplar against which to view cognitive modeling. In circadian research, the models are not proposals regarding the basic architecture of circadian mechanisms; rather, they are used to better understand the functioning of a mechanism whose parts, operations, and organization already have been independently determined. In particular, circadian modelers probe how the mechanism's organized parts and operations are orchestrated in real time to produce dynamic phenomena—what we have called *dynamic mechanistic explanation* (Bechtel & Abrahamsen, in press).

We begin with an overview of mechanistic explanation in general. We then develop the case of circadian rhythm research, where the architecture has been highly constrained by empirical

inquiry into the physical mechanism and modeling is directed to understand the mechanism's dynamics. We do this by examining in turn six different exemplars from the research literature on computational modeling of circadian rhythms. In all of these cases computational modeling was needed to understand the behavior of a complex mechanism involving nonlinearly interacting components. In examining their particulars, though, we draw out six more specific contributions of computational modeling. We then go through these six contributions again, this time presenting for each a cognitive model and querying to what extent it might make the same kind of contribution. This review of models also brings to light certain differences between cognitive scientists and circadian modelers in how they approach computational modeling.

1. Mechanisms and Mechanistic Explanation

Many philosophical presentations of cognitive science (and other sciences) continue to focus on laws as the explanatory vehicle. Laws are commonly construed as universal generalizations that have a modal status—they identify not just what has happened when particular conditions are met, but what must happen under those conditions. But cognitive scientists, and indeed life scientists generally, seldom propose laws. When they do (in psychology, typically referring to them as *effects*), they generally serve not to explain but to characterize the phenomenon to be explained (Cummins, 2000). When they advance explanations, life scientists commonly seek to uncover the mechanism responsible for the phenomenon of interest. Recently, a number of philosophers whose focus has been largely on biology have attempted to characterize what scientists mean by a mechanism and how they go about developing and evaluating mechanistic explanations (Bechtel & Abrahamsen, 2005; Bechtel & Richardson, 1993; Glennan, 1996, 2002; Machamer, Darden, & Craver, 2000; Thagard, 2006). Our own 2005 characterization began:

> A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization.

Discovering the parts and operations of a mechanism requires decomposing it. This typically necessitates experimental techniques since in naturally occurring mechanisms, especially living systems, the parts and operations are so highly integrated that they cannot be identified directly. It is relatively easy to find ways to fracture a system into parts of some sort—the challenge is to identify the *working parts* that perform the operations producing the phenomenon of interest. In the case of the brain, a variety of approaches have been pursued. In the 19th century, the focus was on the sulci and gyri created by the folding of the cerebral cortex, and while these still are used as anatomical landmarks, they are not regarded as working parts. Once it was recognized that cortex comprised individual cells—neurons—neuroanatomists such as Brodmann (1909/1994) used the presence of neurons of specific types and especially differences in the thickness of the layers into which they were organized to differentiate regions in the cerebral cortex. His clear hope was that these areas had functional significance, but he lacked tools for determining this. Refined in later decades using such criteria as neural connectivity and topographical mapping, and studied functionally using such techniques as single-cell recording, it turned out that Brodmann's areas demarcated working parts of the brain so well that they still are in use (Mundale, 1998).

Identifying operations usually involves a very different set of experimental procedures than identifying parts. The goal is to identify operations that do not produce the phenomenon

individually but only in collaboration with other operations performed by different parts of the mechanism (otherwise there is no explanatory gain from decomposing the mechanism). Detecting the effects on overall behavior from experimental manipulations of particular parts (e.g., ablating or stimulating them) often provides suggestive clues, as does recording specific internal effects of altering the inputs to the mechanism. Whatever technique is chosen, proposing operations on the basis of the outcome typically requires elaborate inferential schemes (Bechtel, in press-a) that can lead to blind alleys, overemphasis on particular operations to the exclusion of others, and additional sources of dispute. The challenges in identifying both parts and operations make mechanistic explanation a long and complex endeavor, but in numerous domains of biology well-supported, enduring accounts have eventually been achieved, providing a foundation for more advanced research.

Discussions of mechanistic explanation often allude to the importance of how the components are organized, but this has been the least developed aspect both of philosophical accounts of mechanistic explanation and of mechanistic science itself. Much more attention has been paid to ways of *decomposing* a mechanism into component parts and operations than to ways of *recomposing* them into an appropriately organized system. Generally scientists use the simplest organizational scheme that will serve their immediate purpose. For example, since the 1930s and still today the main backbone of reactions in glycolysis has been represented as a linear sequence (plus side reactions): Glucose $\rightarrow$ G6P $\rightarrow$ F6P and so forth—not unlike a diagram of a simple assembly line. Yet, as biological theorists from Claude Bernard to the present have recognized, there are distinctive modes of organization in organisms that enable them to exhibit such phenomena as maintaining themselves in a non-equilibrium relation to their environment. Recognition first of negative feedback and later of positive feedback and self-organizing cycles have offered biologists a more precise understanding of the key role of organization in living systems (Bechtel, 2006, 2007).

Such modes of organization orchestrate the parts and operations in real time. Thus, our 2005 characterization of mechanism continued as follows:
> The orchestrated functioning of the mechanism is responsible for one or more phenomena.

Though this orchestration often is downplayed as investigators focus on identifying parts and operations, attending to it can reveal complex dynamics, ranging from periodic oscillations to chaos. Different tools than those employed in early investigations of a mechanism are required to pursue its dynamics: the tools of quantitative computational modeling. These tools have a long history and have been employed in a variety of ways, often disconnected from any sort of mechanistic project. For example, the system of interest may not fruitfully be described as a mechanism, or the modeler may prefer to focus on global variables and parameters. When explicitly anchored to a particular mechanistic account, however, computational modeling enables exploring and understanding the dynamics of that mechanism. In short, it offers dynamic mechanistic explanation.

To extend mechanistic explanation to accounts of dynamics, we augment the above characterization of mechanism with the phrase in boldface:
> A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the

mechanism, **manifested in patterns of change over time in properties of its parts and operations,** is responsible for one or more phenomena.

In this paper we focus on the most common way of achieving dynamic mechanistic explanation: computational modeling with differential equations in which time is one of the variables. To pursue this strategy, the modeler selects properties of certain parts or operations of the mechanism that appear to be salient to a particular dynamic phenomenon. These properties are then pulled into a computational model as variables or parameters, thereby anchoring that model to the mechanistic account. The modeler has flexibility in this kind of project; for example, a property initially treated as a parameter may, as the model is further explored, become a variable. These strategies for developing dynamic mechanistic explanations are evident in the modeling of circadian mechanisms to which we now turn.

2. Mechanistic Explanation and Modeling in Circadian Rhythm Research

The ability of organisms to keep track of the time of day, even when deprived of external cues such as exposure to sunlight, has fascinated investigators since ancient times (Androsthenes of Thasus, a captain in Alexander's fleet, recorded the daily movement of the leaves of the tamarind tree, while Hippocrates and Galen both observed how body temperature in patients with fevers varied with time of day). Subsequently, circadian rhythms have been found in a wide variety of living organisms, from cyanobacteria to plants, fungi, and a variety of animals such as Drosophila, mice, and humans. They affect biochemical processes (e.g. protein synthesis), physiological functions (e.g., digestion), behavioral phenomena (e.g., locomotor activity), and cognitive performance (e.g., reaction times). Figuring out how they do so proved challenging. Systematic study of these phenomena—incorporating experimental methods in addition to finer-grained description—began only in the middle of the 20th century and soon yielded a richer characterization. In particular, it was learned that circadian rhythms are endogenously controlled but entrainable by Zeitgebers (environmental cues such as onset of daylight or temperature changes) and are temperature compensated (the rhythms maintain nearly the same periodicity irrespective of temperature).

In the quest for the mechanism underlying circadian rhythms in mammals, Stephan and Zucker (1972) and Moore and Eichler (1972) established a central role for the suprachiasmatic nucleus (SCN), a structure in the anterior hypothalamus consisting of approximately 8,000-10,000 neurons on each side of the brain. Their primary evidence was that lesions to the SCN left the organism arrhythmic. Inouye and Kawamura (1979) found circadian rhythms in the SCN's electrical activity even when it was removed from the organism. The final demonstration of the centrality of the SCN to circadian behavior was provided by Ralph, Foster, Davis, and Menaker (1990). They transplanted the SCN from a mutant hamster (whose circadian period was significantly less than 24 hours) into a non-mutant hamster whose SCN had been lesioned and showed that the recipient had the same abbreviated circadian period as the donor. A few years later Welsh, Logothetis, Meister, & Reppert (1995), using multielectrode arrays to record activity in tissue extracted from the SCN, determined that individual SCN neurons maintained rhythms, albeit with a wide range of periods across neurons. Similar advances were made during this period in locating the central clock in *Drosophila* (fruit flies). This organism initially became the focus of circadian rhythm research because its eclosion (emergence from the pupa) is closely

timed to dawn: Regardless of the time of day when a fruit fly completes its development, it delays eclosing until the subsequent dawn. Moreover, even if the pupae are kept in total darkness, they eclode at what would have been dawn (Pittendrigh, 1954). The search for the responsible mechanism in *Drosophila* identified a small number of lateral neurons as playing the central role in maintaining circadian time.

Having localized the responsible mechanism in two target species, the longer-term project has been to figure out its parts and operations and how they are coordinated so as to generate circadian rhythms. In the remainder of this section we discuss six advances at the intracellular and intercellular levels of investigation that provided the foundation for new computational models. For each advance we highlight one such model and address the contribution it makes to understanding circadian mechanisms and their dynamics.

Identifying and modeling the first clock component: PER

Explaining how neurons maintain rhythms required identifying parts within them that individually perform different operations but jointly generate a 24 hour rhythm. The pioneering research involved inducing mutations chemically in *Drosophila* and screening for mutants whose eclosion manifested aberrant circadian rhythms. Konopka and Benzer (1971) found a genetic locus at which mutations resulted in rhythms with shortened or lengthened periods or in loss of rhythms altogether, and designated the responsible gene *period* (*per*). The cloning of *per* in the mid-1980s by Michael Rosbash and his colleagues made it possible to fill in additional components of the molecular mechanism responsible for circadian rhythms. The expression of *per* results in increased concentrations of *per* mRNA and, in turn, of the protein PERIOD (PER) that is synthesized from *per* mRNA in the cytoplasm.[1] Hardin, Hall, and Rosbash (1990) determined that the concentrations of both *per* mRNA and PER exhibited circadian oscillations, with the peaks and valleys in PER concentrations lagging behind those of *per* mRNA by approximately 8 hours. Moreover, since PER was detected not only in cytoplasm (where it is synthesized) but also in the nucleus, they proposed a mechanism involving a negative feedback loop to explain the circadian oscillations. As illustrated in Figure 1, *per* is transcribed *into per-*mRNA; this macromolecule then is transported to the cytoplasm and translated into a protein, PER. PER subsequently is transported back into the nucleus, where it slows synthesis of additional PER by inhibiting *per* transcription. A way for *per* to be released from inhibition was suggested by Edery, Rutila, and Rosbash (1994). They discovered that the molecular mass of PER changed through the day, indicating that it was undergoing phosphorylation. Since phosphorylation is often a prelude to the breakdown of a molecule, this pointed to a process through which PER could be degraded. With less PER in the nucleus, there would be less (or no) inhibition of *per*.

---

[1] Gene and protein names are commonly abbreviated to three letters. Protein names are written in uppercase (PER) and gene names in italics—lowercase for *Drosophila* (*per*) and first letter capitalized for mammals (*Per* or *mPer1*).
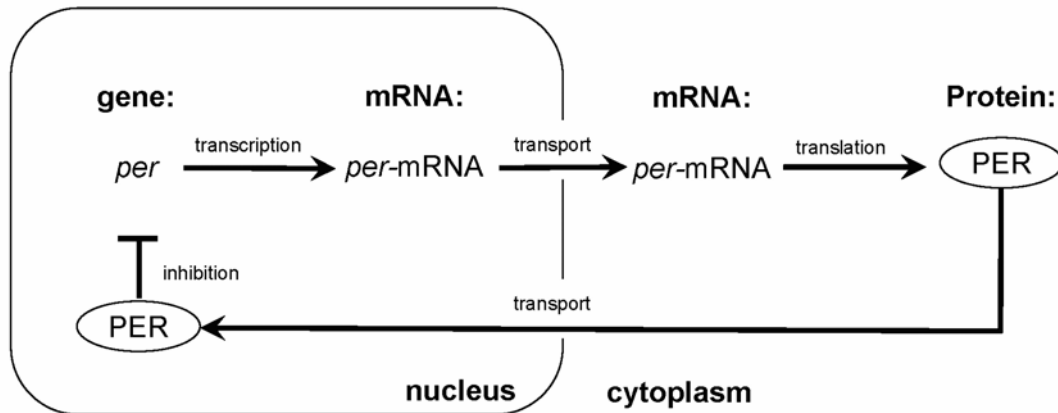
Figure 1. Hardin, Hall, and Rosbash's (1990) proposed feedback mechanism for generating circadian oscillations in PER concentrations.

It is possible to describe verbally how the mechanisms of negative feedback (portrayed in Figure 1) and degradation produce oscillations. Assuming low protein concentrations in both the nucleus and cytoplasm at the start, the gene acts to synthesize more of its protein (by transcription into *per*-mRNA, which must be transported into cytoplasm before it can be translated into PER). Over several hours, PER accumulates in cytoplasm. As its concentration rises, some of the molecules are transported into the nucleus, where by some unspecified mechanism they inhibit the gene that earlier had produced them. As less new PER is synthesized and existing PER degrades, its concentration in cytoplasm now is falling rather than rising. This results in fewer PER molecules entering the nucleus. With those already there degrading, gradually there is less inhibition of *per*, and therefore a swing back to a higher rate of protein synthesis that soon is reflected in rising cytoplasmic PER concentrations.

By rehearsing these steps, one can mentally simulate the generation of a periodic oscillation from the mechanism. However, to demonstrate that a sustained oscillation would result—rather than, for example, a gradual dampening to a steady state—a quantitative model of the process was required. To provide this, Goldbeter (1995) adapted a model that Goodwin (1963) had created for a related purpose: exploring the dynamics of the feedback process in Jacob and Monod's (1961) operon model of gene regulation.[2] Goldbeter's version of the model employed five kinetic equations that yielded predicted rates of change in concentrations of *per* mRNA and PER, based in part on the rates of these key operations:

---

[2] Goodwin was interested in the conditions under which oscillations in protein synthesis would occur. He identified as the key parameter the Hill coefficient n, which also figures in Goldbeter's model (see equation presented below) and represents the number of molecules that must cooperate to achieve inhibition. Using analog simulations, Goodwin had reported oscillatory behavior with values of n as low as 2 or 3, but shortly afterwards Griffith (1968) showed that undamped oscillations would only occur with $n>9$, a value he deemed too high to be realistic. Griffith took this to show that negative feedback involving a single gene product on a gene could never "give rise in practice to undamped oscillations in the concentrations of cellular constituents." Others, including Goldbeter, overcame this obstacle by incorporating nonlinearities elsewhere in the model (e.g., in the terms specifying degradation of various components).

- Transcription of *per* into *per* mRNA in the nucleus (subject to inhibition from nuclear phosphorylated PER) and its transport into cytoplasm
- Degradation of *per* mRNA in cytoplasm
- Synthesis of PER in cytoplasm (rate proportional to the concentration of *per* mRNA)
- Reversible phosphorylation of PER in cytoplasm
- Degradation of phosphorylated PER in cytoplasm
- Reversible transport of phosphorylated PER between cytoplasm and nucleus

Nonlinearity in the equation describing accumulation of cytoplasmic *per* mRNA (*M*) is particularly important to the generation of oscillations in this model:

$$\frac{dM}{dt} = v_s \left(\frac{K_I^n}{K_I^n + P_N^n} - v_m \frac{M}{K_m + M} \right)$$

Here v_S represents the maximum rate for the accumulation of *M*, K_I is a threshold constant for inhibition, P_N is concentration of PER in the nucleus, *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. For some values of its parameters, the system described by the five equations will quickly settle into a steady, non-oscillatory state. By running simulations with various combinations of parameter values, Goldbeter identified values at which the steady state condition gives way to limit cycle oscillations in the concentrations of *per* mRNA and PER of the sort observed in *Drosophila*. When values of *per* mRNA and PER fall on the limit cycle (dark oval in Figure 2), they repeat the same pattern of change indefinitely. If the initial values fall outside the cycle, they follow a transient trajectory (one of the spiraling lines) until they join the limit cycle. Goldbeter focused in particular on a parameter v_d, the maximum rate at which PER is degraded before entering the nucleus, and found that between two critical values, the period of oscillation varied between 19.3 hours and 64 hours. (The precise effect of v_d also depended on the value of other parameters, such as k_s, the rate of protein synthesis). Goldbeter claimed that the ability of his model to generate alterations in period length by varying one specific parameter offered a possible explanation of the mutant forms with lengthened or shortened periods that Konopka had found.
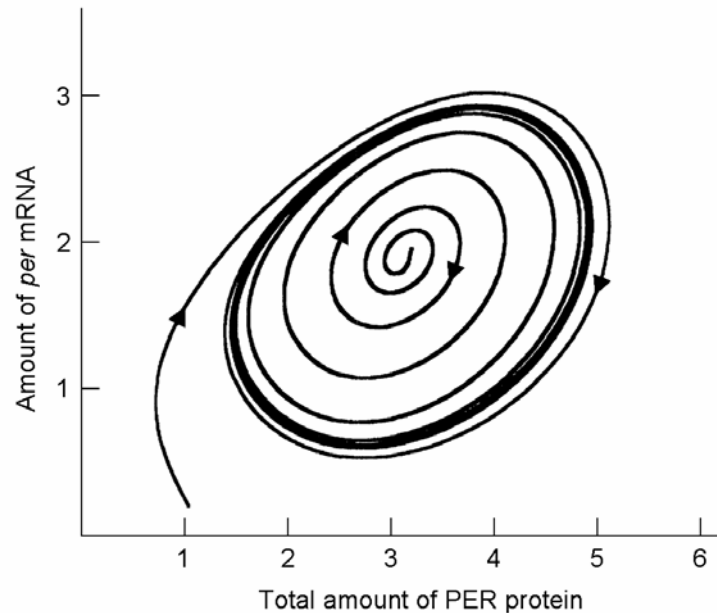
Figure 2. Limit cycle generated by Goldbeter's (1995) model. See text for explanation.

By showing that, with appropriate parameters, a model employing the proposed feedback of PER on *per* translation and transcription could generate circadian oscillations, Goldbeter's research illustrates the first type of contribution computational modeling can make:

> **(1) A model can demonstrate that a mechanism whose parts, operations, and organization have been at least partially identified is able to exhibit the phenomenon of interest.**

Moreover, the model itself shows why computational modeling, not just mental simulation, is required for this—there are many values of the parameters under which the phenomenon does not occur. Mental simulation alone cannot reveal these. Goldbeter's model only generates what Craver (2007) calls a "how possibly" explanation, but when the issue is whether a particular organization of components could realize an effect, that itself is an important contribution to understanding the mechanism.

Adding and modeling a second clock component: TIM

While Goldbeter's model seemed to suggest that those components of the mechanism which he included might be sufficient to explain circadian oscillations, the molecular biologists knew at least one important component was missing. This is because the PER molecule lacks the region (domain) needed to bind to DNA and function as a transcription factor in inhibiting its own transcription. A suggestion as to the missing component was provided by the fact that PER was homologous to two other proteins that were known to be transcription factors (the three were the first known members of the PAS group). The suggestion was that PER might form a compound with one of these (or with a yet unknown transcription factor), and that it would be this conjoined protein that suppressed *per* transcription and translation.

To search for such a transcription factor, Sehgal, Price, Man, and Young (1994) generated *Drosophila* mutants in much the same manner as Konopka. One of their mutants manifested no rhythmic behavior, and they named the responsible gene *timeless* (*tim*). In a further study, these researchers established that in this mutant, PER did not enter the nucleus and was unstable in the cytoplasm (Vosshall, Price, Sehgal, Saez, & Young, 1994). Subsequently they also determined that PER and TIM form a dimer (a compound of two similar units) in cytoplasm and it is this dimer that is transported back into the nucleus. Moreover, they established that a region found on both PER and TIM, which they named the *cytoplasmic localization domain* (CLD), was responsible for preventing either molecule alone from migrating into the nucleus (Gekakis et al., 1995). They suggested that this region was masked in the dimer and that this allowed the dimer to be transported into the nucleus (Saez & Young, 1996).

It turned out, though, that TIM, like PER, lacked a region for binding with DNA. Thus, it could not resolve the original question of how *per* transcription was inhibited. Yet, its discovery pointed to an explanation for another important circadian phenomenon, the entrainment of oscillations by light. Exposure to light in early evening results in a phase delay in circadian behavior, whereas exposure in the pre-dawn hours produces a phase advance. (Increased exposure to light during the day has neither effect.) The fact that TIM is broken down when fruit flies are exposed to light suggested a mechanism for this. In early evening TIM levels are rising and breakdown slows their accumulation, delaying inhibition of *per*. In late evening, in contrast, TIM levels are decreasing and breakdown hastens that process, hastening the release of *per* from inhibition (Hunter-Ensor, Ousley, & Sehgal, 1996).

To determine how TIM might contribute to the working of the circadian mechanism, Goldbeter extended his earlier model in collaboration with Jean-Christophe Leloup (Leloup & Goldbeter, 1998; see also Tyson, Hong, Thron, & Novak, 1999). The extended model required ten differential equations and created limit cycles comparable to those of the PER-only model. By varying parameter values, Leloup and Goldbeter sought to understand the relation between phosphorylation of PER and TIM, the dimerization process, and the degree of cooperativity (number of cooperating molecules) required to enter the nucleus. They showed that with greater cooperativity, the range of other parameters in which oscillations could be maintained was much broader, and that requiring phosphorylation and dimerization both extended these ranges. In exploring these conditions in the model, Leloup and Goldbeter were going beyond simply showing that the mechanism might suffice for generating circadian rhythms; they were beginning to use the model to explore how the mechanism might respond under a variety of conditions. In particular, to explore how light exposure could change the oscillatory phase by degrading TIM, Leloup and Goldbeter focused on v_{dT}, a variable in one of the equations in their model which represents the maximum rate of TIM degradation. To simulate conditions of constant darkness they specified a fixed value of v_{dT}, and to simulate light exposure they doubled that value. Depending on the interval during which v_{dT} was doubled, the model generated phase advances or phase delays comparable to those induced by light exposure in *Drosophila*. This provided supported for the mechanism of entrainment proposed by Hunter-Ensor et al.

One of the benefits of working with a computational model is that modelers can explore a variety of parameter values which, if the equations correctly describe the behavior of components of the mechanism, correspond to changes in the mechanism that may be difficult to generate

experimentally. For example, Leloup and Goldbeter found that some parameter values yielded two stable oscillatory regimes with different periods, a phenomenon known as birhythmicity, and others that produced chaotic oscillations. Although commenting "It probably is too early to speculate on their possible physiological significance, particularly in view of the reduced range of parameter values in which they occur," they nonetheless proposed that this might account for the birhythmicity observed empirically by Pittendrigh (1960) upon changes in environmental conditions. Pittendrigh thought this pointed to multiple oscillators, but Leloup and Goldbeter suggested it may be due to identical oscillators in different cells responding to different initial conditions.[3] (In a further elaboration of the model which we will not discuss here, Leloup & Goldbeter, 1997, showed how it could account for temperature compensation.)

These explorations with the model point to a second contribution of computational modeling of mechanisms:

> **(2) A model provides a means of exploring a much larger space of parameter values than would be feasible experimentally, and thereby of projecting how the actual mechanism would behave under a variety of conditions.**

Modelers, accordingly, often speak of conducting experiments with their models. These experiments, unlike experiments on the physical mechanism itself, do not show that the mechanism in fact could produce these effects. The equations cannot capture all salient aspects of the mechanism and may even misrepresent or leave out critical components. But insofar as the focus in modeling is on the effects of the organization of a system, they do demonstrate that if a particular mode of organization were realized in the actual system, the effects should correspond to those in the model.

Discovering and modeling a second feedback loop

The original mystery of how PER inhibited translation and transcription of *per*, which engendered the search for more clock components, eventually was solved by researchers searching for clock mutants in mice. This was much more challenging than the earlier efforts with *Drosophila* in that mice live longer and produce fewer offspring. Nonetheless, within just two weeks of beginning their attempt, Takahashi, Pinto, & Vitaterna (1994) generated a mutant with a long circadian period that, when homozygotic, resulted in total loss of rhythms. They named the responsible gene *Clock* (for Circadian Locomotor Output Cycles Kaput) and determined that its protein, CLOCK, also oscillated. When the same laboratory succeeded in cloning *Clock* three years later, they predicted, based on the expected amino acid sequence of the protein, that it "encodes a novel member of the bHLH–PAS domain family of transcription factors" (King et al., 1997, p. 645). The PAS domain is the one found in PER, whereas bHLH (basic helix-loop-helix) is a DNA binding domain. In virtue of possessing this domain, CLOCK, unlike PER or TIM, is able to bind to *per* (specifically, to the binding region on *per* denoted E-box promoter). King et al. predicted that a *Clock* homologue existed in *Drosophila*, and indeed Darlington et al. (1998) found it the next year. The *Drosophila* CLOCK protein was also found to have a dimerization partner, named CYCLE despite the fact it doesn't cycle. (BMAL1 is the dimerization partner of CLOCK in mammals.) Darlington et al. therefore proposed that (1) the CLOCK:CYCLE dimer activates the transcription of both *per* and *tim* by binding to their E-

---

[3] For more recent explorations of ways to generate birhythmicity through control of the timing of light exposure in hamsters, see Gorman (2001).

boxes; and (2) when the PER:TIM dimer enters the nucleus, it can inhibit the CLOCK:CYCLE dimer from performing this function—perhaps by removing it from the E-box on the *per* and *tim* genes. As shown schematically in Figure 3, (1) corresponds to subjective nighttime and (2) to subjective daytime.
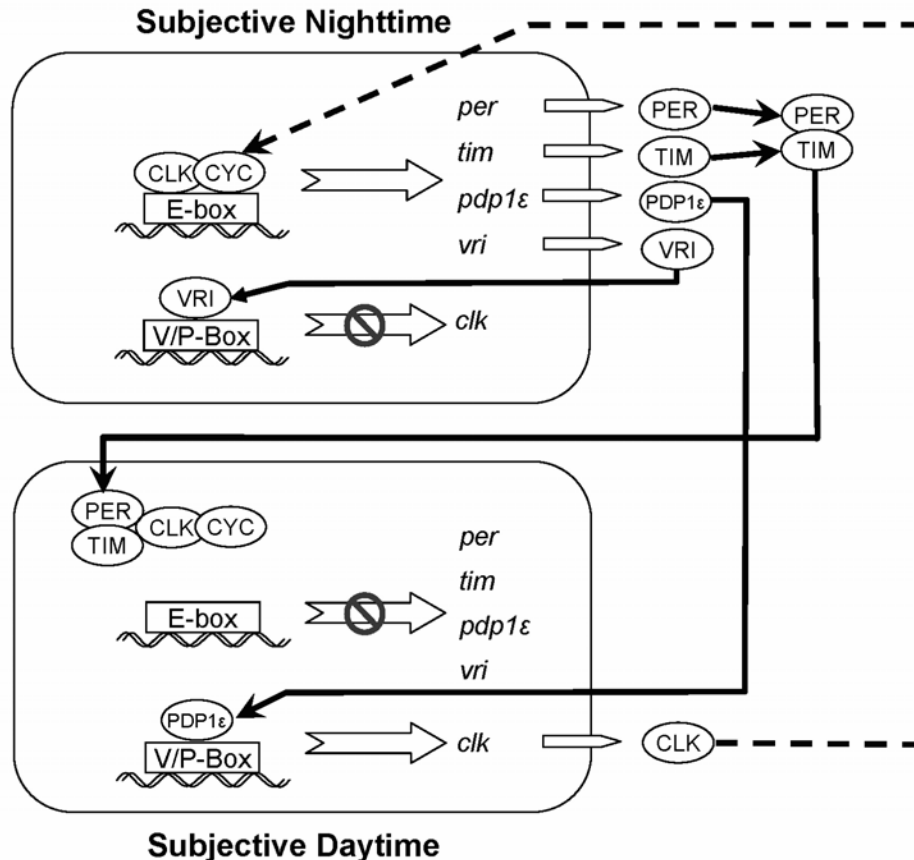

Figure 3. A more complete account of the *Drosophila* oscillator. The large open arrows indicate whether the promoter turns gene expression on or off. The smaller open arrows represent the combined processes of gene transcription in the nucleus, transport to the cytoplasm, and translation in the cytoplasm.

As Figure 3 makes clear, the CLOCK:CYCLE dimer not only excites the transcription of *per* and *tim* but also excites *clock* transcription (via the action of PDP1ε on the V/P box on *clock*). This is a positive feedback loop: an increase in *clock* transcription results in increased PDP1ε production; and the additional PDP1ε binds with the V/P box on *clock*, resulting in yet further increase in *clock* transcription. The intermediate roles of VRILLE and PDP1ε were only discovered by Cyran et al. (2003) after the modeling efforts described below, but the fact that CLOCK feedback positively on its own production was already anticipated. (VRILLE initially inhibits *clock* transcription, but this can be ignored since its effect is soon supplanted by PDP1ε. The overall feedback of CLOCK on its own transcription is positive.)

These discoveries raised the question of how the positive feedback loop (involving CLOCK) related to the negative feedback loop (involving PER:TIM). Did it "cancel out" the negative feedback? Or was it crucial to maintaining oscillations, as Hastings (2000) proposed? To address

these questions, Smolen, Baxter, & Byrne (2001) developed an alternative model that was much simpler than Leloup and Goldbeter's and included only the processes deemed essential to the oscillation. For example, they did not differentiate the roles of PER and TIM, but combined them into one variable. They also did not attempt to model transport between the cytoplasm and the nucleus but instead included a parameter for the delay between changes in concentration of CLOCK and changes in the rate of generation of additional CLOCK and PER/TIM. Their model yielded oscillations in total CLOCK and PER/TIM concentrations with a period of 23.6 hours, a result that was robust over substantial variation in the various parameters. To investigate whether CLOCK was essential to maintaining rhythmicity they fixed the value for CLOCK concentrations in the model. The model continued to generate oscillations, and these remained robust even in the face of substantial changes in the other parameters. They concluded that the positive feedback loop involving CLOCK was not critical to the generation of circadian rhythms; rather, the circadian oscillations were due solely or primarily to the negative feedback loops involving PER and TIM.

This reveals a different function for computational modeling of a mechanism:
> **(3) A model can indicate which parts and operations of a system that have been identified empirically are essential for producing the phenomenon of interest, and therefore should count as components of the responsible mechanism.**

Although it might be possible to conduct manipulations in an experimental system similar to those performed in the model (e.g., by finding a buffer or otherwise stabilizing the concentration of CLOCK), such exploration is much easier to accomplish in a model. It is noteworthy that in this type of modeling researchers need to capture in detail only the parts they think might be relevant and can use a simplified account of other parts of a mechanism. Although there are purposes for which it is important to have a relatively complete model, Smolen et al's work. demonstrates how simplifying the model can also be useful. In a simpler model is it easier to identify which components of the system are sufficient to produce the phenomenon and hence comprise the responsible mechanism, and which are components of the system but not of the responsible mechanism for that particular phenomenon.

The mammalian oscillator and modeling pathologies

After a *Clock* homologue was found in *Drosophila*, mammalian researchers in turn investigated whether mammalian homologues might exist for many of the genes first found in *Drosophila*. Two groups of researchers identified a mammalian homologue to *per* (Sun et al., 1997; Tei et al., 1997) and determined that the resulting protein was 44% identical to the *Drosophila* protein (with many of the differences involving neutral amino acid substitutions). Soon after it was recognized that in fact there are multiple mammalian homologues of *per*, designated *mPer1*, *mPer2* (Albrecht, Sun, Eichele, & Lee, 1997), and *mPer3* (Zylka, Shearman, Weaver, & Reppert, 1998). The search for homologues also revealed some important differences between the *Drosophila* and mammalian clocks. CRYPTOCHROME (CRY) was identified in *Drosophila* as subserving entrainment by light—it is a photoreceptor and, in response to light, induced the degradation of TIM (Stanewsky et al., 1998). In mammals, however, CRY usurps the role of TIM as the primary dimerization partner of PER (Griffin, Staknis, & Weitz, 1999) and melanopsin replaces CRY for entrainment (Hattar et al., 2003). Accompanying these changes were differences in how the clock mechanism works—for example, when CRY plays an

entrainment function, it promotes the breakdown of TIM; in contrast, melanopsin operates by activating production of PER. (For a discussion of how assuming evolutionarily conserved mechanisms facilitated these discoveries, see Bechtel, in press-b.)

Differences such as these led Leloup and Goldbeter (2003, 2004) to model the mammalian mechanism by modifying their *Drosophila* model. Their basic finding was that the mammalian mechanism, like the *Drosophila* mechanism, could produce limit cycle oscillations. The focus on mammals, though, enabled an extension of the goals of modeling. Humans in particular exhibit a variety of circadian pathologies, two of which are delayed sleep phase syndrome (DSPS: attributed to those whose natural pattern is to fall asleep several hours after midnight and wake up in the late morning) and advanced sleep phase syndrome (ASPS: attributed to those who get very sleepy in the early evening and wake up around 3 a.m.). Leloup and Goldbeter found that, by altering the values of parameters characterizing the phosphorylation of PER, they could mimic these syndromes. This comported well with the finding by Toh et al. (2001) that in one family with inherited ASPS, a point mutation had occurred in which a glycine molecule replaced a serine molecule in the casein kinase Iδ/ε (CKIδ/ε) binding domain, which figures in phosphorylation of PER. Subsequently Xu et al. (2005), studying another family with inherited ASPS, found a mutation in the CKIδ gene (in which an alkaline substituted for a threonine). This further supported Leloup and Goldbeter's focus on parameters describing the phosphorylation of PER to account for these pathologies.

An important measure in assessing a mechanistic account is the extent to which it can capture the effects of various kinds of damage to the mechanism. In physiology, this means accounting not only for normal functioning but also for various pathologies. Here is another contribution of computational modeling:
>**(4) A model offers a possible explanation for a disorder when specific alterations to the model produce effects that mimic the disorder.**

In this the modeling goes beyond the empirical research which only showed what was different in the pathological cases, not how the altered factor could produce the symptoms.

Modeling synchronization between cellular oscillators

The research by Welsh, Logothetis, Meister, and Repert (1995) demonstrating that individual SCN neurons dispersed in culture sustained oscillations with a mean period of approximately 24 hours also revealed a large range of variation (21.25 to 26.25 hours) and standard deviation (1.2 hours). In contrast, Herzog, Aton, Numano, Sakaki, and Tei (2004) noted that there is much less variability in measurements of period length for running wheel behavior in mice and established that this was also true for SCN slices in which the spatial relations between neurons were maintained. This indicated that individual neurons were synchronizing with each other. Synchronizing oscillators, though, is a non-trivial problem. The dynamics tend to get very complex, including toroidal oscillations, deterministic chaos, or coexistence of multiple attractors (Grebogi, Ott, & Yorke, 1987). Several groups of researchers have chosen computational modeling as an especially suitable strategy for understanding how synchronization is achieved among neurons in the SCN.

Aton, Colwell, Harmar, Waschek and Herzog (2005) proposed that vasoactive intestinal polypeptide (VIP) was a likely synchronizing agent. To investigate how release of VIP might play a role in synchronization, Gonze, Bernard, Waltermann, Kramer, and Herzel (2005) developed a model in which a population of oscillators stood in for the neurons of the SCN. They modeled each oscillator in a population using a variant of the oscillator Goodwin had proposed for the operon by viewing it as describing the generation and degradation of a single clock protein (e.g., PER). To the basic equations they add one describing change in concentration of VIP in individual oscillators (one term describes its rate of generation as proportional to the current concentration of the clock protein and a second describes its rate of degradation as proportional to its own current concentration) and another equation 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. (periods that were highly variable across oscillators), but when it was set to 0.5, the oscillators synchronized. Analyzing the case of just two oscillators, Gonze et al. were able to suggest that the oscillations of individual neurons in the SCN might dampen without VIP and that sustained oscillatory behavior is due to VIP.

Gonze et al. relied on mean VIP levels rather than modeling diffusion, but To, Henson, Herzog, and Doyle (2007) offered a more complex model that included details of diffusion as well as the pathway by which VIP affects *Per* transcription. They also randomly perturbed a parameter specifying basal transcription of *Per* mRNA such that only about 40% of the model SCN cells sustained oscillation in the absence of VIP. (Aton et al. had found that in the mouse SCN only 30% of cells appear capable of sustaining oscillations on their own.) They succeeded in replicating Aton et al.'s empirical findings—when VIP was present the cells synchronized, but when VIP was removed approximately 60% became arrhythmic and the rest desynchronized. They also replicated a finding that in constant light, individual SCN cells continue to oscillate but are desynchronized (Ohta, Yamazaki, & McMahon, 2005).

Although synchronization between oscillators is known to sometimes give rise to complex dynamics very different from those of circadian rhythms, these two simulations suggest that relatively straightforward processes of peptide release and uptake would produce the sort of synchronization found in SCN neurons. In this context computational modeling serves another function:

> **(5) A model can reveal conditions under which independent mechanisms, each with its own internally determined dynamical (oscillatory) behavior, can be coupled so as to exhibit coordinated collective behavior.**

Moreover, the models suggested rather direct ways to account for the entrainment properties of light on populations of oscillators and explain phenomena associated with different light regimes.

Desynchronization between oscillators and modeling jet lag

Although we emphasized just above the achievement of synchronization between oscillators, there are conditions that result in desynchronization. One of the most familiar is jet lag, in which clock-controlled functions (most notably sleep, but also various metabolic and cognitive

activities) are disrupted for many days after traveling multiple time zones, especially in the eastward direction. The problem is not that the central clock is slow to adjust to entrainment signals in the new environment, but that we have multiple oscillators; those that become desynchronized require considerable time to resynchronize.

It is now recognized that normal circadian behavior requires not only synchronization among a homogeneous population of cells, but also coupling between distinct populations of cells so as to maintain normal phase offsets between them. The SCN has long been recognized as containing two populations of cells—a ventrolateral or core region (these are the cells that produce VIP) and a dorsomedial or shell region (these cells generate vasopressin, Moore, Speh, & Leak, 2002).[4] Oscillations are synchronized within each region, but out of phase between regions: PER reaches maximum concentrations nearly an hour earlier in the shell than in the core (Nakamura, Yamazaki, Takasu, Mishima, & Block, 2005). Moreover, most organs of the body (the liver, heart, and brain have been most studied) have cells whose oscillation is coordinated by the SCN with varying phase delays. It was long assumed that these peripheral oscillators were unable to sustain oscillations (rather, oscillations dampened) without input from the SCN; accordingly, they were regarded as slaves to the SCN. More recently evidence suggests that they maintain oscillations, but lose synchrony. Davidson, Yamazaki, and Menaker (2004) appeal to this evidence to argue that the SCN is better viewed as a conductor that synchronizes the oscillation of peripheral oscillators than as a ringmaster that spurs them into action.

To explore the effects of jet lag, Nakamura et al. subjected rats to either to a six hour delay (equivalent of traveling westward, achieved by delaying turning off lights by 6 hours and creating an 18 hour day) and a six hour advance (corresponding to traveling eastward, achieved by turning lights on 6 hours early, creating a 6 hour night). Although neurons in the SCN core (the only ones to receive entrainment signals from the eye) adjusted relatively quickly, the synchronization between different oscillators was highly disrupted. This was especially true with a six hour advance (corresponding to eastward travel). After one day the core SCN had shifted significantly more than the shell SCN, resulting in an inversion of the normal order of the cycles. In fact, core SCN had advanced more than 9 hours, a significant overshoot of the 6 hour change in the light schedule. By day three the overshoot in the core had reduced whereas the shell had advanced nearly 6 hours, resulting in virtually no phase differences between them. It was only on day 6 that the normal phase relation, with the shell leading the core by nearly an hour, was restored, and even then both had advanced more the expected 6 hours.

To explore such effects between the SCN core and shell, and also between those and peripheral oscillators, Leise and Siegelmann (2006) developed a model. They used just two equations to describe the oscillation of two state variables, *X* and *Y* (whose biochemical interpretation they leave unspecified), in each oscillator:

---

[4] One of the intriguing features of the core region is the presence of a dense collection of calbindin D28K (CalB) cells (Silver, LeSauter, Tresco, & Lehman, 1996). When these cells are destroyed but the remainder of the SCN neurons are not altered, overt physiological and behavioral rhythms are eliminated (LeSauter & Silver, 1999). Yet, these calbindin cells do not themselves oscillate either in neuronal firing or in *Per* gene expression. Rather, PER1 and PER2 are synthesized in them in response to photic stimulation (Hamada, LeSauter, Venuti, & Silver, 2001; see also Shirakawa, Honma, & Honma, 2001). Antle, Foley, Foley, and Silver (2003) advanced a proposal that these calbindin cells function to gate signals to other SCN cells and explored how this could work in a computational model.

$$\frac{dX}{dt} = \frac{r_X}{1+Y^2} - q_X X$$

$$\frac{dY}{dt} = r_Y X^2(t - t_{lag}) - q_Y Y$$

Initially X peaked before Y such that plotting values of X against those of Y reveals a limit cycle. To model entrainment by light, they reduced the value of r_Y to indicate when light was available, and restored it to indicate night. Then to model a time zone change, they shifted the values of time t for which r_Y was reduced. They found that after a simulated six-hour advance or ten-hour delay, the oscillator exhibited a transient two-hour overshoot that gradually reduced until the appropriate limit cycle was again attained.

In the next step Leise and Siegelmann modeled connections among six oscillators representing the two populations of SCN oscillators and four populations of peripheral oscillators (about which different assumptions were made concerning whether the oscillations were damped and the phase lag from the core SCN). The oscillator representing the core SCN was connected to one representing the shell SCN, and that in turn was connected to the four peripheral oscillators. Coupling between core and shell oscillators was accomplished by increasing the value of r_Y, and thereby concentrations of Y, in the shell oscillator when the value of X in the core oscillator exceeded a threshold. Coupling between shell and peripheral oscillators was accomplished by increasing the value of r_X in the peripheral oscillator when the value of X (or, in other cases, Y) exceeded a threshold. To simulate the effects of a six-hour advance, they temporarily reduced the value r_Y in the equations describing the core SCN oscillator and it advanced rapidly and initially overshot. Initially the shell SCN oscillator did not change, but it subsequently advanced and itself overshot the target. After six days the peripheral oscillator appear to have regained the original phase relation to the core oscillators, but then overshot further and required seven more days to again achieve the original phase relation.

The peripheral oscillators also required approximately two weeks to restore their phase relation with the two populations of SCN oscillators, although different patterns of restoration resulted from different assumptions built into the model. When the choice of parameter values rendered the peripheral oscillators as damped oscillators, they recovered their phase relation relatively quickly, but when they were self-sustained this took longer and required a strong degree of coupling. Under some conditions, the peripheral oscillators restored phase relations by delaying fifteen hours rather than advancing six hours. Leise and Siegelmann went on to explore how more gradual phase advances affected the synchronization of clocks. When the phase advance was spread over four days or when the period of darkness was extended two hours during the first two cycles, resynchronization was much more orderly and achieved much more quickly than with the simple six-hour phase advance.

At present there are a great many unknowns involving both the nature of the coupling between SCN and peripheral oscillators and how different transition regimes affect the restoration of synchrony. In this situation, models can serve a different function:

(6) A model provides a means of exploring the space of possibilities for altering and restoring relations between multiple mechanisms (e.g., synchrony or phase delay between oscillators).

Given the current uncertainty as to whether peripheral oscillators are damped oscillators or merely unsynchronized when they receive no input from the SCN, modeling phase advances (and delays) with both kinds of oscillators can reveal the empirical signatures that each may leave (e.g., delaying fifteen hours rather than advancing six) and so guide empirical attempts to answer this question. Using models to further explore how systems of coupled oscillators may respond to different procedures for changing periods may also guide further empirical search for strategies for coping with jet lag and related phenomena. In these cases, the modeling not only contributes to evaluating explanations of accounts built up from empirical inquiry, but potentially can guide empirical research as well.

Summary comments on circadian models

Offering circadian rhythms as an illustrative case, we have described six contributions that computational modeling can make to the understanding of mechanisms. The models we discussed were constructed only after laboratory-based researchers had learned a good deal about the parts, operations, and organization of the biological mechanisms responsible for various circadian phenomena. Computational modelers took advantage of these basic mechanistic explanations but confirmed and extended them by deploying quantitative tools for understanding system dynamics. Often these tools are applied abstractly, either because the system cannot fruitfully be described as a mechanism or because the modeler prefers to focus on global variables and parameters. When substantially anchored to a particular mechanistic account, however, computational modeling helps to understand and explore the dynamics of that mechanism—yielding what we have called dynamic mechanistic explanation.

4. Modeling and Mechanistic Explanation in Cognitive Science

The relation between modeling and mechanistic explanation is very different in cognitive science than in circadian rhythm research and many other areas of biology. For the most part, empirical research in cognitive science has not revealed the representations or other component parts of cognitive mechanisms or their operations (Bechtel, 2008). Instead, cognitive scientists generally posit these components in their computational models and then do empirical research to demonstrate that they can produce the phenomena of interest. If they can, the modeler claims that the cognitive mechanism is like that specified in the computational model. (If not, the modeler usually revises the model and tests it again.). This can be seen by considering more specifically the two forms of computational modeling that, until recently, have dominated cognitive science—symbolic modeling and neural network modeling.

Symbolic modelers posit mental representations in which discrete symbols are composed into strings, trees, or other symbolic structures that are crucial parts of the cognitive system. They presume that cognitive activity involves rules that generate and operate on these representations. Among the many variations on this basic architecture originating in the early decades of cognitive science are Chomsky's tree structures, rewrite rules and transformational rules and artificial intelligence systems operating on representations inspired by the predicate calculus or semantic networks, scripts, and so forth. One of the most enduring is Newell and Simon's (1972) production system architecture. Its key components are a working memory in which symbolic structures are stored and a set of conditional statements called *production rules* of the form: "if

X, do *Y*." When the condition is satisfied (e.g., the fact that the box is not on the table is in working memory) the production rule will *fire*, directing the system to perform action *Y*, (e.g., sending an instruction to put the box on the table, or modifying the contents of working memory). A given production system model (e.g., EPAM, SOAR, ACT*) is treated as a hypothesis about the kinds of representations and operations implemented in the brain. Initially the evidence for such a model was limited to behavioral data (e.g., patterns of errors or reaction times or, more recently, eye movements), as these were the most relevant available types of data and cognitive psychologists had become skilled at procuring and interpreting them. Beginning in the mid-1980s, the introduction of PET and then fMRI neuroimaging technologies, as well as advances in ERP, provided non-invasive means of securing neural evidence regarding cognitive operations. Initially this gave rise to a rapidly expanding field of cognitive neuroscience. Increasingly, though, mainstream cognitive scientists such as Anderson (2007) have sought to link the operations posited in computational models to specific brain areas by examining neural activation as people performed tasks designed to call upon particular hypothesized cognitive operations.

Artificial neural network accounts posit a different architecture, one in which cognitive activity is presumed to involve propagation of activation across networks of neuron-like processing units (hence, these also are known as connectionist networks). In *parallel distributed processing* (PDP) networks (Rumelhart & McClelland, 1986b), it is patterns of activity over sets of units that count as representations (sometimes referred to as *sub-symbolic*), but there is also a class of *localist* networks in which each unit serves a representational function on its own (and hence might be regarded as a discrete symbol). In both, the operations that propagate activation are specified by equations applying to each individual unit (summing inputs from weighted connections into a unit, calculating an activation level for the unit, and sending output to other units). Below we will present selected neural network models in greater detail, but here we note that they are designed to execute transformations of input representations to output representations appropriately for a given cognitive task—that is, to account for phenomena—not to simulate the activity of known neural pathways (Rolls & Treves, 1998, constitute an important exception).

Despite their differences, symbolic and neural network models share certain points of contrast with circadian models. Circadian modelers begin with an existing mechanistic account and then undertake computational modeling to explore its dynamics and, simply, whether it works. The computational model is a system of differential equations whose variables correspond to selected properties of the parts and operations of the target mechanism. The characterization of the whole mechanism, in detail, is a prior task achieved by other methods (the conduct and interpretation of laboratory research). But for cognitive scientists the computational model, with its hypothesized representations and operations, is the only available account of the mechanism. Depending on the type of model (they are quite diverse in cognitive science), it may also capture dynamics qualitatively or quantitatively. Cognitive operations and their dynamics are assumed to be mirrored in the model, but the only empirical check on this is the extent to which the model as a whole approximates the behavior of humans performing the target task. Thus, much more is

asked of computational models in cognitive science than in circadian research, but they benefit less from empirical constraint.[5]

Despite substantial differences in how computational modeling is carried out, cognitive modelers can make some of the same kinds of contributions as circadian modelers. We will illustrate this by briefly presenting for each of the six contributions already discussed one comparison case from cognitive science. All involve connectionist neural network models, and were discussed at greater length in Bechtel and Abrahamsen (2002, chapters 8 and 10). However, the same points could have been made using symbolic models of the 1980s and 1990s as well as most contemporary cognitive science models. These have in common that they are not based on empirically grounded knowledge of the target mechanism and hence have a more hypothetical status than the corresponding circadian models. In each case we will note how this limits the evaluation or interpretation of the model.

>   **(1) A model can demonstrate that a mechanism whose parts, operations, and organization have been at least partially identified is able to exhibit the phenomenon of interest.**

The models proposed in cognitive science often serve a similar, limited objective: showing that a mechanism of the sort hypothesized in the model would be adequate to produce one or more phenomena of interest. Success in meeting this objective provides weak evidence for the model as a whole (it is shown to be plausible) and indirect evidence that the posited operations correspond to those of the actual mechanism. For example, Rumelhart and McClelland (1986a) wished to demonstrate that a single connectionist network was adequate to account for children's U-shaped pattern in learning to produce the past-tense of both regular and irregular verbs (e.g., the correct form *went* gets replaced by the overregularized form *goed*, eventually corrected to *went*). As a network that could learn, each time a verb was presented to its input units it applied (a) its operations for propagating activation to its output units and (b) a learning rule, which assessed how well the output pattern corresponded to the target past-tense form and adjusted the connection weights towards a slightly better response next time. The same verbs were presented hundreds or thousands of times, and discrepancies from the correct past-tense forms could be examined at different points in learning and at the end. Since the model was not grounded in research on the actual mechanism, it could be evaluated only by the similarity of these discrepancies to those exhibited by children—that is, its ability to exhibit relevant phenomena. To their credit, Rumelhart and McClelland did not limit themselves to the particularly striking

---

[5] The fact that the computational models of cognition employ operations that mirror those taken to be involved in the cognitive mechanism helps explain the fact that the models themselves are often taken to perform cognitive activities. If, for example, operations over representations are hypothesized as enabling the cognitive agent to make intelligent decisions, and the model employs similar operations over its representations, then the model itself, if successful, is capable of making intelligent decisions. In contrast, one would not regard a computational model of a circadian mechanism as itself generating circadian rhythms (or a computational model of a hurricane as itself being a hurricane). This is because the model of circadian rhythms (or of the hurricane) does not involve operations corresponding to those in the mechanism being modeled. Rather, variables in the model reflect salient properties of the parts and operations in the mechanism and the equations capture how values of these variables change over time. The computational model, if successful, tracks how the values of these variables change through time. The computational model describes aspects of the working of the mechanism; it does not mirror its operation.

phenomenon of a U-shaped acquisition curve across the learning process; they identified a number of other phenomena in the existing empirical literature, some rather subtle ones involving subtypes of phonologically similar verbs, and for the most part found that their model came impressively close to replicating them.

Critics too focused on the fit between the model's behavior and empirical phenomena, but brought to the table additional phenomena and more stringent criteria with the result that they judged the model's fit inadequate (Pinker & Prince, 1988). Favoring a more traditional symbolic explanation, they argued for a dual mechanism account: rules for regular verbs and memory look-up for irregular verbs. The result was a vigorous debate in which subsequent modelers (e.g., Plunkett & Marchman, 1993) attempted to overcome the shortcomings identified by critics, critics formulated additional objections, and so forth over multiple rounds. This controversy illustrates one of the limitations of modeling in the absence of knowledge of the parts, operations, and organization of the actual mechanism (which for cognitive phenomena is, at some level of abstraction, a brain mechanism). In the absence of this source of constraint on the numerous ways one or more competing models might be modified to accommodate additional phenomena, the ensuing conflicts often are irresolvable. The models are underdetermined, and the modelers themselves perhaps overly determined to win an unwinnable competition.

> **(2) A model provides a means of exploring a much larger space of parameter values than would be feasible experimentally, and thereby of projecting how the actual mechanism would behave under a variety of conditions.**

This second contribution captures the fact that by altering the values of parameters, investigators can perform experiments on models that are difficult or impossible to perform in the real world. The benefits are available even to cognitive scientists limited to hypothetical models, especially when they target not just the components of a mechanism but also its dynamics. Beer (1995, 2000), for example, used the walking behavior of six-legged insects as a domain for investigating the extent to which connectionist networks could get outcomes similar to those of actual neural control circuits. He was guided by general knowledge about how insects walk and the circuits involved, but his models were intended as an abstraction of the broad class of actual circuits. Each "leg" was controlled by a network of five fully interconnected units, three of which (akin to motor neurons) also sent outputs to effectors in the leg. All five units also had the potential to receive input from sensors in the leg, and there were interconnections among the six controller networks (one for each leg) so as to coordinate their behavior. The weights on the various connections were obtained via a genetic algorithm (that is, rather than using a learning algorithm to train up the weights in a single network, Beer used selection over multiple generations to evolve networks with weights well-adapted to the task). The resulting networks reproduced the tripod gait characteristic of actual fast-moving six-legged insects. Most relevant here, Beer conducted experiments in which he varied sensory feedback. Consistent input during simulated evolution yielded networks that reproduced fine-tuned walking with the tripod gait—but if sensory input was withheld during testing they could not walk. In contrast, networks evolved under sensory deprivation walked with a stereotyped gait regardless of whether sensory input was provided during testing. Finally, networks for which the sensory inputs were sometimes on and sometimes off during evolution produced a fine-tuned tripod gait when tested with sensory input and the stereotyped gait otherwise.

Like the experiments on circadian models, the experiments Beer conducted in evolving his model controller networks were illuminating in exploring parameter space (he regarded environmental input as a parameter in his broader framework). He also made advantageous use of dynamical systems theory (DST) in further interpreting the findings: since his networks had connections in both directions between most units, they were interactive (rather than feedforward) networks and therefore formed attractor basins for which DST provides revealing analytic tools. As a motor control model grounded in part on neural research, but abstracting from it, Beers' model is less hypothetical than the others considered here. Direct inferences would still be risky, but his results suggest that blocking or degrading sensory pathways to the extent possible might be an intriguing avenue of research on actual neural control systems. And were more realistic versions of his model developed, experiments on them could more plausibly substitute for research in those regions of parameter space that are intractable for actual systems.

> **(3) A model can indicate which parts and operations of a system that have been identified empirically are essential for producing the phenomenon of interest, and therefore should count as components of the responsible mechanism.**

A model of memory consolidation offered by McClelland, McNaughton, and O'Reilly (1995) sought to make a similar contribution. Their model was situated at the intersection of cognitive science and cognitive neuroscience, in that they started with a fairly well-supported claim that consolidation involved two different kinds of brain structure—hippocampus and neocortex—and their modeling was directed to explaining **why**. It was not obvious why consolidation of long-term memories in neocortex should take months, nor why the hippocampal system (which specializes in quickly creating intermediate-term, contextualized memories) should be necessary to that process. McClelland et al. proposed that a three-layer feedforward network that learns via *backpropagation* provided a reasonably close model of the type of memory mechanism instantiated in neocortex. (This type of network has a middle layer of *hidden units*; trained repeatedly on a set of input-output patterns, its learning rule gradually modifies the weights on the connections from input units to hidden units and from hidden units to output units until the output patterns closely approximate the target output patterns for each input.) A well-known problem with this architecture is that it exhibits the phenomenon of catastrophic interference, in which learning a new set of input-output pairs severely compromises performance on those learned previously. This can be avoided if the training regimen interleaves new items with reinstated old items. McClelland et al. proposed that the hippocampus served as trainer, interleaving new items with old items that it "replayed" to neocortex. Due to memory decay in hippocampus, the older the item the less often it would be replayed. They provided this kind of training to a feedforward network serving as a computational model of the neocortical memory mechanism. (They provided no model of the hippocampus—that was a later project—but only the interleaved output it was presumed to send to neocortex.) When McClelland et al. stopped the training (corresponding to lesioning the hippocampal system), they found that the feedforward network (neocortex) performed worst on items that had the fewest consolidation trials (i.e., it showed the graded retrograde amnesia that is a signature of hippocampal damage). As intriguing as are the results of this modeling effort, the ability to apply them to the actual interaction of the hippocampus and neocortex is limited by the fact that we do not yet know whether hippocampus provides appropriate interleaved training to neocortex or that neocortex

functions like a feedforward network subject to catastrophic interference. The modeling does advance hypotheses that, once empirical researchers find a means of testing them, could make important contributions to our understanding of mammalian memory mechanisms.

(4) A model offers a possible explanation for a disorder when specific alterations to the model produce effects that mimic the disorder.

A similar contribution has been offered in certain connectionist modeling projects, except that a connectionist investigator typically alters a particular set of connections rather than resetting one or a few parameters in a set of equations. To cite one especially illuminating example, Hinton and Shallice (1991) designed a network to model the role of semantics in reading aloud. Positing a path from orthographic representations (spellings) to semantic representations (meanings) to phonological representations (for reading aloud), they implemented the first half of the path and assumed that the second half was unproblematic. Specifically, they started with a feedforward network in which orthographic units sent activation to semantic units. The challenge was that similar input patterns tend to activate similar output patterns, but orthographically similar words can have quite different meanings. They therefore amended the architecture by adding interconnections between semantic units as well as a set of "clean-up units" that received inputs from the semantic units and sent outputs back to them. This yielded an interactive network that could develop attractor basins for different meanings. Thus, although two words with similar orthographic representations would otherwise tend to activate similar semantic representations, the attractor basins served to pull them apart and the network performed well. Hinton and Shallice then explored what happened when different sets of connections were "lesioned" by removing a subset. They found three types of errors: visual errors (reading *CAT* as "mat"), semantic errors (reading *MOUSE* as "rat"), and mixed errors (reading *CAT* as "rat"), in somewhat different proportions for different lesions (e.g., more semantic errors with lesions higher in the network). Their proposed explanation was that lesions restructured the attractor basins. Such errors are characteristic of the pathological condition known as *deep dyslexia*. Their model made available a novel explanation of this disorder that was quite appealing, especially in that reading researchers previously had been baffled by the co-occurrence of semantic and visual errors and a higher than expected rate of mixed errors. Despite genuine success, the same caveat applies to this case as to those already discussed: because the computational model was not anchored to an account of the actual neural mechanism responsible for reading aloud (either intact or altered so as to produce deep dyslexia), any inferences that the model's architecture and specific operations explain the ability or the disorder are supported only indirectly. In the absence of a better-grounded explanation one might choose to rely upon that of Hinton and Shallice, but only tentatively and as a guide to future research.

(5) A model can reveal conditions under which independent mechanisms, each with its own internally determined dynamical (oscillatory) behavior, can be coupled so as to exhibit coordinated collective behavior.
and
(6) A model provides a means of exploring the space of possibilities for altering and restoring relations between multiple mechanisms (e.g., synchrony or phase delay between oscillators).

These last two contributions of computational modeling address the coordination, or the disruption of coordination, between component mechanisms whose activities are determined primarily by their internal constitution. In cognitive science, most modelers treat components of their models as merely responding to inputs. Some, though, treat them as oscillators that are intrinsically active. For example, to account for the spontaneous way in which people switch between alternative interpretations of ambiguous figures such as the Necker cube, van Leeuwen and colleagues developed a *coupled map lattice* model comprising 2500 (50 x 50) oscillator units (van Leeuwen, Steyvers, & Nooter, 1997; van Leeuwen, Verver, & Brinkers, 2000). They used the logistic function.

$$x_{t+1} = A\, x_t\,(1 - x_t)$$

as an activation function for each unit, where x_t is the net input to that unit at time *t*. Depending on the value of the parameter A, a unit on its own can exhibit periodic or chaotic oscillations (or not oscillate). However, each unit was coupled to four other units (its neighbors in the lattice) at a strength determined by the value of a coupling parameter C. With appropriate values of C relative to A, synchronization was achieved within small clusters of units.

To simulate shifting interpretations of ambiguous figures, some elaborations were required. The modelers redefined A as a variable that was sensitive to input and replaced C with individually adaptive weights plus a global parameter C_{max} that could bias the model towards more or less stability. Presented with a grid of equally spaced dots, vertical neighbors in the lattice would temporally synchronize their oscillations (i.e., the model would see the grid as columns of dots). At some point units would desynchronize, and the next synchronization might involve horizontal rather than vertical neighbors (i.e., the model would see the grid as rows of dots). The model exhibited *metastable synchronization* as the interpretations irregularly alternated. This is an impressive demonstration of how one computational model that focuses on coupling relations between oscillators can mimic a particular perceptual phenomenon. Like the second and fourth cognitive science models introduced above, it specifies both a mechanism and its performance dynamics. As well, like the first and fourth models, it simulates learning dynamics. But like all four of the above cognitive science models, it was developed in the absence of information that the relevant parts and operations in the brain realize such a mechanism. At best it suggests how and why the actual perceptual mechanism behaves this way if it in fact employs oscillatory components organized and orchestrated like those in the van Leeuwen model.

5. Conclusion

In summary, computational modelers who focus on cognitive capacities use many of the same computational tools and seek to make many of the same contributions as those focused on circadian rhythms. Both can offer dynamic mechanistic models as explanations for one or more phenomena. However, the models we offered in illustration of these similarities also displayed an important difference: whether component parts and operations are posited in the model or discovered through empirical inquiry prior to modeling.

Cognitive modelers begin by designing a computational model that hypothesizes a mechanism with specified parts, operations, and organization, and then try to show that such a model can mimic the cognitive phenomena of interest. If successful, they offer it as a model of the actual mechanism. Some of these go beyond basic mechanistic modeling to include dynamic

phenomena among the explanatory targets; in particular, those we cited addressed dynamics at a timescale of either real-time performance or ontogenesis. While this is a welcome advance in cognitive modeling, it does not change the status of these models as de novo proposals for the architecture of cognitive mechanisms. Empirical evidence can rule out specific models, but usually cannot decide between competing architectures.

In contrast, computational models of circadian mechanisms, like those advanced in many other sciences, elucidate the functioning of a mechanism whose parts, operations, and organization already have been identified by empirical research in which instrumentation, experimentation, and laboratory methods were tools of discovery. The variables in such a model capture salient properties of the parts, operations, and organization described in the mechanistic account, and the model's performance helps scientists assess whether such a mechanism could realize the dynamic phenomena of interest. This is particularly important when the operations are nonlinear and the organization is sufficiently complex that mental simulation of the mechanism's functioning would be inadequate.

Computational modeling in circadian rhythm research hence provides a different type of exemplar against which to view cognitive modeling. It suggests a path for cognitive modeling that can attend to dynamics without ignoring mechanism: equations in which the variables include properties of parts and operations as well as time. It also highlights the value of computational modeling that is anchored in an empirically-derived mechanistic account. This kind of dynamic mechanistic explanation will not come easily in cognitive science, in that it must target extremely complex mechanisms. However, the necessary empirical and conceptual tools are increasingly available, as is the will to use them in this way.

References

Albrecht, U., Sun, Z. S., Eichele, G., & Lee, C. C. (1997). A differential response of two putative mammalian circadian regulators, *mper1* and *mper2*, to light. *Cell, 91*(7), 1055-1064.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford: Oxford University Press.

Antle, M. C., Foley, D. K., Foley, N. C., & Silver, R. (2003). Gates and oscillators: A network model of the brain clock. *Journal of Biological Rhythms, 18*(4), 339-350.

Aton, S. J., Colwell, C. S., Harmar, A. J., Waschek, J., & Herzog, E. D. (2005). Vasoactive intestinal polypeptide mediates circadian rhythmicity and synchrony in mammalian clock neurons. *Nature Neuroscience, 8*, 476-483.

Bechtel, W. (2006). *Discovering cell mechanisms: The creation of modern cell biology*. Cambridge: Cambridge University Press.

Bechtel, W. (2007). Organization and biological mechanisms: Organized to maintain autonomy. In F. C. Boogerd, F. J. Bruggeman, J.-H. Hofmeyr & H. V. Westerhoff (Eds.), *Systems biology: Philosophical Foundations*. New York: Elsevier.

Bechtel, W. (2008). Mechanisms in cognitive psychology: What are the operations? *Philosophy of Science, 75*.

Bechtel, W. (in press-a). The epistemology of evidence in cognitive neuroscience. In R. Skipper, C. Allen, R. A. Ankeny, C. F. Craver, L. Darden, G. Mikkelson & R. C. Richardson (Eds.), *Philosophy and the Life Sciences: A Reader*. Cambridge, MA: MIT Press.

Bechtel, W. (in press-b). Generalization and discovery through conserved mechanisms: Cross species research on circadian oscillators. *Philosophy of Science*.

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. (in press). Complex biological mechanisms: Cyclic, oscillatory, and autonomous. In C. A. Hooker (Ed.), *Philosophy of complex systems. Handbook of the philosophy of science, Volume 10*. New York: Elsevier.

Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton, NJ: Princeton University Press.

Beer, R. D. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence, 72*, 173-215.

Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences, 4*, 91-99.

Brodmann, K. (1909/1994). *Vergleichende Lokalisationslehre der Grosshirnrinde* (L. J. Garvey, Trans.). Leipzig: J. A. Barth.

Craver, C. (2007). *Explaining the brain: What a science of the mind-brain could be*. New York: Oxford University Press.

Cummins, R. (2000). "How does it work?" versus "what are the laws?": Two conceptions of psychological explanation. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 117-144). Cambridge, MA: MIT Press.

Cyran, S. A., Buchsbaum, A. M., Reddy, K. L., Lin, M.-C., Glossop, N. R. J., Hardin, P. E., et al. (2003). vrille, Pdp1, and dClock Form a Second Feedback Loop in the Drosophila Circadian Clock. *Cell, 112*, 329-341.

Darlington, T. K., Wager-Smith, K., Ceriani, M. F., Staknis, D., Gekakis, N., Steeves, T. D., et al. (1998). Closing the circadian loop: CLOCK-induced transcription of its own inhibitors *per* and *tim*. *Science, 280*(5369), 1599-1603.

Davidson, A. J., Yamazaki, S., & Menaker, M. (2004). SCN: Ringmaster of the circadian circus or conductor of the circadian orchestra? In D. J. Chadwick & J. A. Goode (Eds.), *Molecular clocks and light signalling* (pp. 110-125). Chichester, UK: John Wiley.

Edery, I., Rutila, J. E., & Rosbash, M. (1994). Phase shifting of the circadian clock by induction of the *Drosophila period* protein. *Science, 263*, 237-240.

Gekakis, N., Saez, L., Delahaye-Brown, A.-M., Myers, M. P., Sehgal, A., Young, M. W., et al. (1995). Isolation of *timeless* by PER Protein Interaction: Defective Interaction Between *timeless* Protein and Long-Period Mutant PER[L]. *Science, 270*(5237), 811-815.

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., & Herzel, H. (2005). Spontaneous synchronization of coupled circadian oscillators. *Biophysical Journal, 89*(1), 120-129.

Goodwin, B. C. (1963). *Temporal organization in cells; a dynamic theory of cellular control processes*. London: Academic.

Gorman, M. R. (2001). Exotic photoperiods induce and entrain split circadian activity rhythms in hamsters. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology, 187*, 793-800.

Grebogi, C., Ott, E., & Yorke, J. A. (1987). Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics. *Science, 238*(4827), 632-638.

Griffin, E. A., Jr., Staknis, D., & Weitz, C. J. (1999). Light-Independent Role of CRY1 and CRY2 in the Mammalian Circadian Clock. *Science, 286*(5440), 768-771.

Griffith, J. S. (1968). Mathematics of cellular control processes I. Negative feedback to one gene. *Journal of Theoretical Biology, 20*(2), 202-208.

Hamada, T., LeSauter, J., Venuti, J. M., & Silver, R. (2001). Expression of *Period* genes: Rhythmic and nonrhythmic compartments of the suprachiasmatic nucleus pacemaker. *Journal of Neuroscience, 21*(19), 7742-7750.

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.

Hastings, M. H. (2000). Circadian clockwork: two loops are better than one. *Nat Rev Neurosci, 1*(2), 143-146.

Hattar, S., Lucas, R. J., Mrosovsky, N., Thompson, S., Douglas, R. H., Hankins, M. W., et al. (2003). Melanopsin and rod-cone photoreceptive systems account for all major accessory visual functions in mice. *Nature, 424*(6944), 75-81.

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.

Hinton, G. E., & Shallice, T. (1991). Lesioning a connectionist network: Investigations of acquired dyslexia. *Psychological Review, 98*, 74-95.

Hunter-Ensor, M., Ousley, A., & Sehgal, A. (1996). Regulation of the *Drosophila* protein Timeless suggests a mechanism for resetting the circadian clock by light. *Cell, 84*(5), 677-685.

Inouye, S.-I. T., & Kawamura, H. (1979). Persistence of circadian rhythmicity in a mammalian hypothalamic "island" containing the suprachiasmatic nucleus. *Proceedings of the National Academy of Sciences (USA), 76*, 5962-5966.

Jacob, F., & Monod, J. (1961). Genetic regulatory systems in the synthesis of proteins. *Journal of Molecular Biology, 3*, 318-356.

King, D. P., Zhao, Y., Sangoram, A. M., Wilsbacher, L. D., Tanaka, M., Antoch, M. P., et al. (1997). Positional Cloning of the Mouse Circadian *Clock* Gene. *Cell, 89*(4), 641-653.

Konopka, R. J., & Benzer, S. (1971). Clock mutants of *Drosophila melanogaster. Proceedings of the National Academy of Sciences (USA), 89*, 2112-2116.

Leise, T., & Siegelmann, H. (2006). Dynamics of a multistage circadian system. *Journal of Biological Rhythms, 21*(4), 314-323.

Leloup, J.-C., & Goldbeter, A. (1997). Temperature compensation of circadian rhythms: Control of the period in a model for circadian oscillations of the PER protein in *Drosophila. Chronobiology International, 14*, 511-520.

Leloup, J.-C., & Goldbeter, A. (1998). A model for circadian rhythms in *Drosophila* incorporating the formation of a complex between the PER and TIM proteins. *Journal of Biological Rhythms, 13*(1), 70-87.

Leloup, J.-C., & Goldbeter, A. (2003). Toward a detailed computational model for the mammalian circadian clock. *Proceedings of the National Academy of Sciences, 100*(12), 7051-7056.

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.

LeSauter, J., & Silver, R. (1999). Localization of a suprachiasmatic nucleus subregion regulating locomotor rhythmicity. *Journal of Neuroscience, 19*(13), 5574-5585.

Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1-25.

McClelland, J. L., McNaughton, B., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419-457.

Moore, R. Y., & Eichler, V. B. (1972). Loss of a circadian adrenal corticosterone rhythm following suprachiasmatic lesions in the rat. *Brain Research, 42*, 201-206.

Moore, R. Y., Speh, J. C., & Leak, R. K. (2002). Suprachiasmatic nucleus organization. *Cell and Tissue Research, 309*, 89-98.

Mundale, J. (1998). Brain mapping. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science*. Oxford: Basil Blackwell.

Nakamura, W., Yamazaki, S., Takasu, N. N., Mishima, K., & Block, G. D. (2005). Differential response of *Period 1* expression within the suprachiasmatic nucleus. *Journal of Neuroscience, 25*(23), 5481-5487.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ohta, H., Yamazaki, S., & McMahon, D. G. (2005). Constant light desynchronizes mammalian clock neurons. *Nature Neuroscience, 8*, 267-269.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28*, 73-193.

Pittendrigh, C. S. (1954). On temperature independence in the clock-system controlling emergence time in *Drosophila*. *Proceedings of the National Academy of Sciences (USA), 40*, 1018-1029.

Pittendrigh, C. S. (1960). Circadian rhythms and the circadian organization of living systems. *Cold Spring Harbor Symposia on Quantitative Biology, 25*, 159-184.

Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition, 48*, 21-69.

Ralph, M. R., Foster, R. G., Davis, F. C., & Menaker, M. (1990). Transplanted suprachiasmatic nucleus determines circadian period. *Science, 247*(4945), 975-978.

Rolls, E. T., & Treves, A. (1998). *Neural networks and brain function*. Oxford: Oxford University Press.

Rumelhart, D. E., & McClelland, J. L. (1986a). On learning the past tenses of English verbs. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2. Psychological and biological models*. Cambridge, MA: MIT Press.

Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986b). *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1. Foundations*. Cambridge, MA: MIT Press.

Saez, L., & Young, M. W. (1996). Regulation of nuclear entry of the *Drosophila*-clock proteins Period and Timeless. *Neuron, 17*, 979-990.

Sehgal, A., Price, J. L., Man, B., & Young, M. W. (1994). Loss of circadian behavioral rhythms and *per* RNA oscillations in the *Drosophila* mutant *timeless*. *Science, 263*, 1603-1606.

Shirakawa, T., Honma, S., & Honma, K.-i. (2001). Multiple oscillators in the suprachiasmatic nucleus. *Chronobiology International, 18*(3), 371 - 387.

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.

Smolen, P., Baxter, D. A., & Byrne, J. H. (2001). Modeling circadian oscillations with interlocking positive and negative feedback loops. *Journal of Neuroscience, 21*(17), 6644-6656.

Stanewsky, R., Kaneko, M., Emery, P., Beretta, B., Wager-Smith, K., Kay, S. A., et al. (1998). The *cry^b* mutation identifies cryptochrome as a circadian photoreceptor in *Drosophila*. *Cell, 95*(5), 681-692.

Stephan, F. K., & Zucker, I. (1972). Circadian rhythms in drinking behavior and locomotor activity of rats Are eliminated by hypothalamic lesions. *Proceedings of the National Academy of Sciences (USA), 69*, 1583-1586.

Sun, Z. S., Albrecht, U., Zhuchenko, O., Bailey, J., Eichele, G., & Lee, C. C. (1997). RIGUI, a putative mammalian ortholog of the Drosophila *period* gene. *Cell, 90*(6), 1003-1011.

Takahashi, J. S., Pinto, L. H., & Vitaterna, M. H. (1994). Forward and reverse genetic approaches to behavior in the mouse. *Science, 264*(5166), 1724-1733.

Tei, H., Okamura, H., Shigeyoshi, Y., Fukuhara, C., Ozawa, R., Hirose, M., et al. (1997). Circadian oscillation of a mammalian homologue of the *Drosophila period* gene. *Nature, 389*(6650), 512-516.

Thagard, P. (2006). *Hot thought: Mechanisms and applications of emotional cognition*. Cambridge, MA: MIT Press.

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.

Toh, K. L., Jones, C. R., He, Y., Eide, E. J., Hinz, W. A., Virshup, D. M., et al. (2001). An hPer2 Phosphorylation Site Mutation in Familial Advanced Sleep Phase Syndrome. *Science, 291*(5506), 1040-1043.

Tyson, J. J., Hong, C. I., Thron, C. D., & Novak, B. (1999). A Simple Model of Circadian Rhythms Based on Dimerization and Proteolysis of PER and TIM. *Biophys. J., 77*(5), 2411-2417.

van Leeuwen, C., Steyvers, M., & Nooter, M. (1997). Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. *Journal of Mathematical Psychology, 41*, 319-344.

van Leeuwen, C., Verver, S., & Brinkers, M. (2000). Visual illusions and outline-invariance in nonstationary activity patterns. *Connection Science, 12*, 279-298.

Vosshall, L. B., Price, J. L., Sehgal, A., Saez, L., & Young, M. W. (1994). Block in nuclear localization of *period* protein by a second clock mutation, *timeless*. *Science, 263*(5153), 1606-1609.

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.

Xu, Y., Padiath, Q. S., Shapiro, R. E., Jones, C. R., Wu, S. C., Saigoh, N., et al. (2005). Functional consequences of a CKIδ mutation causing familial advanced sleep phase syndrome. *Nature, 434*(7033), 640-644.

Zylka, M. J., Shearman, L. P., Weaver, D. R., & Reppert, S. M. (1998). Three *period* homologs in mammals: Differential light responses in the suprachiasmatic circadian clock and oscillating transcripts outside of brain. *Neuron, 20*(6), 1103-1110.