

Some Virtues of Modeling With Both Hands

William Bechtel

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

Webb distinguishes two endeavors she calls *animal modeling* and *animat modeling* and advocates for the former. I share her preference and point to additional virtues of modeling actual biological mechanisms (animal modeling). As Webb argues, animat modeling should be regarded as modeling of specific, but made-up, biological mechanisms. I contend that modeling made-up mechanisms in situations in which we have some knowledge of the actual mechanisms involved is modeling with one hand—the good one—tied behind one’s back.¹ The hand that is used in animat modeling is constructing and evaluating models by whether they behave in the right way—do they exhibit the particular phenomenon one is trying to understand? The good hand that is disavowed seeks to use evidence about the mechanism employed in real living systems both for inspiration in designing the model and for evaluating the model. Denying oneself use of one’s good hand both limits one’s access to valuable evidence for evaluating a model and denies oneself access to a potent discovery strategy.

Webb draws attention to one reason to employ the good hand—if models are to be relevant to biology (and not just characterize hypothetical mechanisms), then the component parts and operations specified in the model must in some way map onto those in actual biological organisms. Especially if one accepts the possibility of multiple realizations, then if one only uses behavior to evaluate the model one may well have described an alternative realization than that found in real organisms. To determine that one has modeled the actual realization, it is necessary to compare the proposed mechanism with the actual mechanism—does it

consist of the same parts, operations, and organization. I have argued elsewhere (Bechtel & Mundale, 1999) that multiple realization is highly suspect if we demand that alternative realizations be compared at the same fine-grain of behavioral detail as employed in claiming that the realizations are actually different. If there is only one realization once we are rigorous in evaluating behavioral equivalence, then models of made-up mechanisms that meet the behavioral test will also correspond to models of actual systems. In fact, however, researchers typically work with a much coarser-grained analysis of the phenomenon to be explained, and in such circumstances the plausibility of mechanisms with very different design generating the same phenomena increases significantly. If the goal is nonetheless to produce models that correspond to actual mechanisms, the sooner we release the good hand and start to use it, the sooner we will arrive at realistic models.

Many animat modelers, however, reject the assumption that the goal is to produce models that correspond to specific actual mechanisms. They advocate instead the importance of idealized or general models—models that represent the breadth of biological possibility not the specific realizations studied in actual biology.² Webb correctly notes that in any modeling of a mechanism, specific assumptions are made about the mechanism. She then questions what makes a made-up mechanism more general than an actual mechanism. This is an important issue and raises a broader question about the nature of generalization in the life sciences.

In philosophy, generalization is usually assumed to take the form of universally quantified statements

Correspondence to: William Bechtel, Department of Philosophy and Interdisciplinary Programs in Cognitive Science and Science Studies, University of California, San Diego, 9500 Gilman Drive, California 92093. *E-mail:* bill@mechanism.ucsd.edu.

such as figure in scientific laws: Newton's gravitational law specifies the attractive force between any two masses. Several philosophers, however, have noted the paucity of laws in biology; explanation in biology instead takes the form of identifying and characterizing the mechanism responsible for a given phenomenon. A mechanism consists of parts that perform operations which are organized to realize the phenomenon of interest (for a discussion of mechanistic explanation in biology, see Bechtel, 2006; Bechtel & Richardson, 1993; Machamer, Darden, & Craver, 2000). The parts, operations, and organization are specific to a given mechanism. Moreover, accounts of mechanisms are commonly developed by investigating particular model organisms (often bred so as to produce individuals with limited variability). How do accounts of mechanisms generalize when they are grounded in specific details about mechanisms found in particular organisms? One key is what biologists speak of as *conserved mechanisms*. These are mechanisms whose parts, operations, and organization have been preserved through evolutionary change. Sometimes conservation seems to involve preservation with almost no change, as is the case with hemoglobin. But more often conserved mechanisms exhibit substantial change. Although many of the genes and proteins in the mechanism responsible for circadian rhythms in mammals are homologs of those found in insects, there have also been large changes. Homologs of given genes perform different operations in the two classes and there has been substantial reorganization of the whole mechanism (Bechtel, in press). Generalization in such circumstances does not involve universal generalizations, but identification of similarities (while also noting differences) between mechanisms produced by a process of descent with variation.

The significance of generalizations developed from constructing animat models is circumscribed because the mechanisms presented in these models are not conserved from extant biological mechanisms, but invented. Accordingly, we cannot situate them in the descent relations of actual organisms. In contrast, animal modeling such as Webb's benefits from being grounded in what is known of the actual mechanism found in crickets. It is true that Webb's models, like all models, abstract from many details of actual crickets, both those known and those not yet known. A modeler selects specific variables, representing particular properties of the parts, operations, and organiza-

tion of a mechanism, to include in the equations of the model. How is this departure from the details of the actual mechanism substantially different than inventing mechanisms that exhibit the relevant phenomena? In both cases modelers depart from the details of extant mechanisms. But in animal modeling the empirical research on the mechanism anchors the endeavor. If the actual mechanism turns out to be substantially different from the abstracted representation invoked in the model, it is the model that must undergo revision.

Abrahamsen and I argue for conclusions similar to Webb's in a paper comparing recent efforts at computational modeling of circadian rhythms with computational modeling as generally pursued in cognitive science (Bechtel & Abrahamsen, in press). In cognitive science, we argue that researchers only specify the architecture of the responsible mechanism in the context of modeling whereas in circadian rhythm modeling empirical research on the mechanism itself provides the architecture of the mechanism. Some of the circadian models we discuss (Goldbeter, 1995; Leloup & Goldbeter, 1998) attempt to represent all details of the underlying mechanism that have been discovered through empirical research. One goal of such modeling is to show that the mechanism could account for the known phenomena. Given the non-linear nature of the interactions of components of the mechanism, this is not a trivial result. Even in these efforts, modeling requires supplying values for a host of parameters; these may be assessed for their biological plausibility, but they cannot simply be determined through empirical investigation. But often the goal is also to explore how modifications to the mechanism (represented in the model by varying parameter values or altering the architecture of the model) would affect its performance. Such efforts serve to determine both how robust the mechanism is and whether modifications can produce interesting phenomena such as behavior corresponding to known pathologies (Leloup & Goldbeter, 2004). It is common to refer to experimenting on a model to determine how it will respond when parameters or parts of it are altered. In many cases, the goal of experiments on models is to explore possibilities that cannot be investigated empirically (sometimes because of technical limitations). Such experiments are still tied to the actual mechanisms being modeled—if it subsequently proved possible to manipulate the actual mechanism in the same manner as the model and it behaved differently, that would count against the model.

Yet another reason to focus modeling efforts on known biological mechanisms is that real biological mechanisms are often rather different than animat modelers might guess. Animat modelers are free to design mechanisms without considerations of phylogeny or ontogeny. The demand to evolve from existing organisms places enormous constraints on the types of mechanisms that now occur—all intermediate steps in the evolution of extant organisms must themselves have been viable. Moreover, mechanisms in adult organisms are typically modifications and refinements of those produced earlier in development, often serving different functions. The result is that often biological mechanisms generate their functions in ways a designer would find quite surprising. For example, rather than using clock mechanisms in individual cells that are highly reliable, animals employ highly variable intracellular clocks that are then synchronized to maintain reliable time. Moreover, even though mammals have a central clock that plays a critical role in synchronization, the basic clock mechanism is found in cells distributed throughout the body and there is suggestive evidence that these distributed clocks maintain oscillations (Welsh, Yoo, Liu, Takahashi, & Kay, 2004). Further, these distributed clocks may modulate the behavior of the central clock. The result is complex dynamical synchronization of many semi-autonomous oscillators. Although at present we are mostly aware of the negative side effects of such a baroque arrangement (viz. jetlag), it may also have advantageous consequences. In any case, it is not a design that an animat modeler is likely to hit upon, but an intriguing feature of actual biology that can productively occupy animal modelers (for such modeling, see Leise & Siegelmann, 2006). Just as truth is sometimes stranger than fiction, actual biological mechanisms may be more interesting than those conceived by modelers. To discover them, modelers need to be free to employ both hands.

Notes

- 1 The phrase is borrowed from Horgan (1997, p. 17), who uses it to characterize connectionist modeling without syntactically structured representations.
- 2 There is a different function that made-up models are claimed to serve—exploring the space of possible mechanisms that might explain a given phenomenon. This may be especially useful when little or nothing is known about the responsible mechanism and researchers are in quest of ideas

about what sort of mechanism to consider. An example is Hinton and Nowlan's (1987) exploration of how learning might affect evolution. While it is certainly plausible that such modeling may facilitate discovery of mechanisms actually realized in organisms, it would be a valuable contribution to investigate how often they have done so.

References

- Bechtel, W. (2006). *Discovering cell mechanisms: The creation of modern cell biology*. Cambridge: Cambridge University Press.
- Bechtel, W. (in press). Generalization and discovery through conserved mechanisms: Cross species research on circadian oscillators. *Philosophy of Science*.
- Bechtel, W., & Abrahamsen, A. (in press). Computational modeling and mechanistic explanation: Understanding circadian rhythms. *Studies in History and Philosophy of Science Part A*.
- Bechtel, W., & Mundale, J. (1999). Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science*, 66, 175–207.
- Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton, NJ: Princeton University Press.
- 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.
- Hinton, G. E., & Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems*, 1, 495–502.
- Horgan, T. (1997). Connectionism and the philosophical foundations of cognitive science. *Metaphilosophy*, 28, 1–30.
- Leise, T., & Siegelmann, H. (2006). Dynamics of a multistage circadian system. *Journal of Biological Rhythms*, 21(4), 314–323.
- 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. (2004). Modeling the mammalian circadian clock: Sensitivity analysis and multiplicity of oscillatory mechanisms. *Journal of Theoretical Biology*, 230(4), 541–562.
- Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67, 1–25.
- Welsh, D. K., Yoo, S.-H., Liu, A. C., Takahashi, J. S., & Kay, S. A. (2004). Bioluminescence imaging of individual fibroblasts reveals persistent, independently phased circadian rhythms of clock gene expression. *Current Biology*, 14(24), 2289–2295.