12 Top-Down Causation in Biology and Neuroscience

Control Hierarchies

William Bechtel

1. Introduction

The notion of top-down causation has been fraught with controversy. Much of this turns on the notion of levels employed. What is it for one entity or causal process to be located at a higher level than another? In the context of biology and neuroscience, an important sense of level arises in the context of control—a controller is at a higher level than the system it controls, and if something is controlling the controller it is at a yet higher level.[1] Thus, transcription factors are at a higher level of control than the genes whose expression they regulate, and neurons are at a higher level of control than muscles and other cells. The circadian clock is at a higher level than the transcription factors whose expression it regulates, and regions of cortex are at a higher level than sub-cortical areas they regulate. My goal is to unpack the notion of top-down causation required to understand the operation of control hierarchies that figure prominently in biology and neuroscience.

Control is exercised on a controlled system. A controlled system consists of a set of processes that causally interact and together bring about some effect. Human-made machines are exemplars of such controlled systems—an automobile consists of a number of parts that perform various different operations that together result in locomotion. In the context of biology and neuroscience, controlled systems (as well as controllers) are commonly referred to as *mechanisms*. In the recent literature on mechanistic explanation, mechanisms have been identified as entities or parts performing activities or operations organized so as to bring about a phenomenon (Machamer, Darden and Craver 2000; Bechtel and Abrahamsen 2005). For example, the heart circulates blood (the phenomenon) as a result of consisting of chambers in which muscles (parts) contract (operation) and valves (parts) limit flow to one direction (operation) in an organized and orchestrated manner.

Although not generally emphasized in philosophical accounts of mechanism, machines as well as mechanisms can be viewed as systems performing

work by constraining the flow of Gibbs free energy[2] (e.g., a pipe channels the free energy of water flowing downhill to move an another object). This work is often done in the service of a larger system of which the machine or mechanism is a part. For work to be performed in a manner that is useful to the larger system, control is needed. This requires that some of the constraints in the machine or mechanism be modifiable; control is exercised by altering these constraints, thereby redirecting the flow of free energy. In the case of a machine such as an automobile, the driver exercises control by, for example, pressing on the accelerator pedal. In traditional engines, there is a linkage from the accelerator pedal to the butterfly valve on the carburetor. The more the valve (the constraint) is pushed open, the more air, along with fuel, enters the combustion chambers of the engine. As a result of the increase of fuel and air, the combustion exerts more force, speeding up the engine's operation. In a living cell, control is also exercised by altering constraints. An enzyme constrains a biochemical reaction and changing the concentration of the enzyme alters the rate of a reaction. The concentration is increased by an activator binding to the promoter site of a gene, allowing more transcription of that gene. Likewise, in a multi-cell organism, for control to be exercised there must be constraints that can be altered. To increase the flow of blood, the contraction of muscles in the various chambers must be increased. This is accomplished through the release of neurotransmitters that bind to receptors in the muscle cell, permitting the formation of cross-bridges between actin and myosin.

To provide a foundation for discussing control of mechanisms, I begin in section 2 by advancing a perspective that situates mechanisms as modules in networks whose endogenous function is modulated by activity elsewhere in the network. In section 3, I turn to human-made machines to introduce a basic mode of control realized by negative feedback. In section 4, I turn back to organisms and discuss why control is even more fundamental in understanding biological mechanisms than in the case of human-made machines and in section 5 consider cases in which feedback provides the needed control. Negative feedback not only is employed directly to control biological mechanisms but also, as I discuss in section 6, is a means of generating oscillations that facilitate controlling at what time a mechanism is operative. In section 7, I turn to neural control, emphasizing its importance in providing hierarchies of control in multicellular organisms whose component cells and mechanisms are endogenously active. I then conclude by emphasizing that top-down causation, as exhibited in the hierarchical control of biological mechanisms, is a fundamental feature of biological and neural systems. Such top-down control doesn't pose any fundamental mysteries since the control mechanisms as well as the controlled mechanisms are all constructed by ordinary mechanisms within the organism.

2. Mechanisms as Modules in Networks

Together with Craver (Craver and Bechtel 2007), I have previously invoked the mechanistic framework in discussing top-down causation. Our concern was to make sense of the idea that changes in a whole mechanism causally affect it components while avoiding the concerns raised by Kim's (1998) exclusion argument. Kim contended that lower-level causal processes, such as those between parts of the mechanism, suffice and preempt any explanatory role for higher-level causes such as the whole mechanism. We argued for a view in which causation should properly be understood as an intra-level relation and that relations between levels should be understood in terms of the constitution relation between a mechanism and its parts. Although we explicitly defended causal interactions between higher-level entities, a shortcoming of our presentation is that it suggested that all causality is in fact at the lower level. In our main examples of a mechanism (a higher-level entity) having an effect on another mechanism (higher-level entity), the effect consisted of altering one or a few parts of the mechanism. Another shortcoming is that we did not explicate the notion of constitution other than saying that a mechanism is constituted by its parts. This leaves the challenge of specifying when a group of entities constitutes a mechanism.

The standard approach to identifying mechanisms has been to include all entities directly involved in the production of the phenomenon. While this generally sufficed in mechanistic biology in the twentieth century, with the development of high-throughput experimental techniques, biologists are discovering vast numbers of additional entities that affect the phenomena for which explanations were sought. Moreover, many of these components are also components of other mechanisms, making it challenging to identify the boundaries of mechanisms. Instead of starting with a phenomenon and identifying the responsible mechanism, systems biologists are increasingly representing the components of whole organisms (typically cells) in networks such as protein-protein interaction networks and gene-regulatory networks. In network representations, nodes stand for entities and edges indicate interactions between entities. The highly integrated character of biological systems is reflected in the fact that these networks are generally shown to be small-world networks (Watts and Strogratz 1998)—networks in which, by traversing a small number of edges, one can pass from a selected node to any other despite most nodes having only a small number of edges to other nodes. The short path length between any two needs seems to jeopardize the ability to identify distinct mechanisms. However, another strategy has provided a way to identify mechanisms even in highly interconnected system. Biological networks typically exhibit high clustering—nodes that are much more densely connected to their neighbors than to nodes elsewhere. Such clustered nodes are characterized as modules and often when nodes

in a network are annotated using labels from ontologies such as the Gene Ontology (Ashburner et al. 2000; Gene Ontology Consortium 2015), modules turn out to correspond reasonably closely to classically characterized biochemical pathways or cell mechanisms (for examples, see Bandyopadhyay et al. 2008; Bandyopadhyay et al. 2010).

In cells, both the small-world character and the occurrence of modules are enhanced by the fact that the number of edges from a node (referred to as its degree) is not distributed randomly. Most nodes have few edges, but a few have a very large number (Barabási and Bonabeau 2003). These are referred to as *hubs—provincial hubs* if they are primarily connected to other nodes of a module and *connector hubs* if they have mostly long-distance connections. Figure 12.1 is a toy example of a network that exhibits modules and both types of hubs.

A network representation such as in Figure 12.1 provides a basis for explicating top-down relations within a mechanistic framework. Modules typically correspond to mechanisms—they consist of the entities (nodes) most of whose operations affect other nodes in the module (reflected in the edges between nodes). The highly interconnected nature of modules often results in endogenous determined dynamical activity within the modules. When the module receives an external input, this endogenous activity pattern is altered, but the behavior is still largely accounted for in terms of the

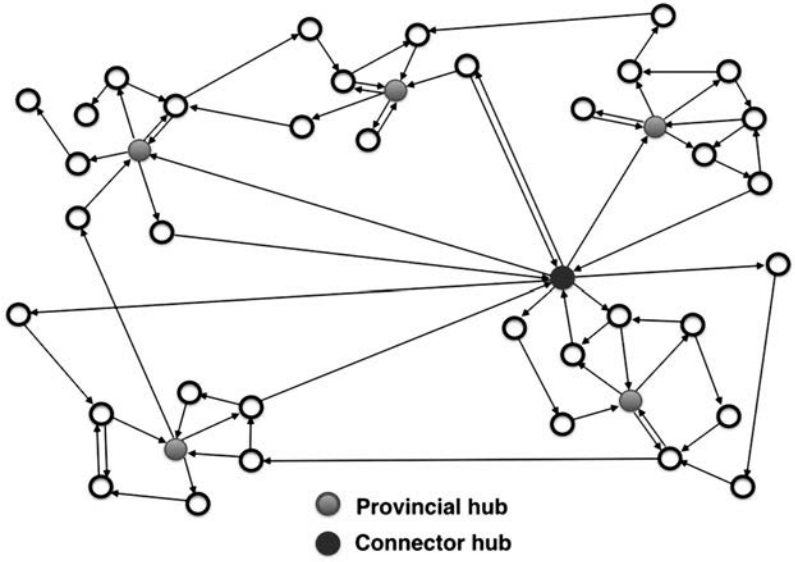

Figure 12.1 A toy example of a network in which there are several modules, each involving a highly connected provincial hub. The network also exhibits small-world properties, as there is a reasonably short path from any node to any other node. This is partly facilitated by a connector hub.

connections (edges) within the module. To understand the behavior of the whole network, we need to recognize both the endogenous activity within the modules and how this activity is affect by inputs from elsewhere in the system (network).

A noteworthy feature of Figure 12.1 is that it does not explicitly distinguish levels; it does, however, support differentiating connections (hence, activity) within and between modules. Misleading, it does suggest that all the entities are at a common lowest level. This is, however, just an artifact of what entities are shown as nodes. If one had reason to consider the inner workings of what is treated as a node, one could expand it into a set of nodes and edges. For each edge connecting to the original node one would have to specify which inner node it connects to (or include multiple edges linking to different internal nodes). Likewise, if the inner organization of modules were no longer of interest, one could replace it with one node and reconnect edges going to inner nodes to the node for the whole module. One should not regard the collection of nodes in a network diagram as at a level but restrict the notion of change of levels to situations in which one identifies parts as constituents of highly interconnected modules (Bechtel (in press)).

While the graph representation provides a foundation for unpacking claims about top-down causation, one shortcoming is that it does not differentiate the kinds of causal interactions edges represent. In particular, it does not distinguish between inputs to and outputs from a controlled mechanism and a controller operating on it. The inputs to the controlled mechanism include the material and energy that figure in producing the output. The controller alters a constraint within the mechanism, often by sending a signal to the entity that provides the modifiable constraint. (The controller performs work, but that work usually requires much less energy and what it does is alter the constraint.) Such a distinction between the inputs to the controlled system and the control processes that modify it is needed if we are to understand the type of top-down causation found in control hierarchies in biology and neuroscience. In the next section, I illustrate control in human made machines.

3. Negative Feedback Control in Human-Made Machines

At least since the time of the ancient Greeks, humans have been building machines to assist in performing work. Many simple machines such as the lever or the screw, constrain and distribute an externally supplied source of Gibbs free energy to effect change (e.g., a human presses down on one side of a lever or turns a screw). More complex machines put together several simple machines; the joint effect of these coordinated machines is still to constrain the release of free energy to perform the desired work. The familiar wing corkscrew, for example, connects two levers to a screw via a set of gears forming a rack and pinion. Together, these constrain the energy

that the user applies so that the screw pulls the cork out of the bottle. More recently, humans have applied other sources of energy (e.g., water flow or electric current) in the machines they build, but these also work by constraining the flow of Gibbs free energy in a manner appropriate for performing the task at hand.

When machines produce the desired work by sequentially executing different operations, control may serve to switch the machine from performing one operation to performing another. In some cases, the human user supplies the control; in the case of the corkscrew, the human initiates and terminates the application of force to the levers. But designers have frequently designed machines in which control is needed but it is impossible or impractical for humans to exercise it. Negative feedback is perhaps the most widely employed design principle for control. Negative feedback employs a second mechanism that responds to a value of a varying property of the controlled system or its output by altering one or more constraints in the controlled system. Negative feedback control appears to have been first employed by Ktesibios in the second century BCE in constructing a water clock. In a water clock, time is registered by the height to which water has risen in a vesicle. For such a clock to keep time reliably, water must be supplied at a constant rate; however, most water sources at the time would not supply a constant input. Ktesibios assured a constant input by inserting a second vesicle between the source and the main vesicle. He maintained the water in this second vesicle at a constant height by employing a float valve that plugged the input pipe except when the water in this vesicle dropped below its target level. Then it would let water in until the target was again achieved. Because water in this vesicle was maintained at a (nearly) constant height, the flow into the main vesicle was at a constant rate. The second vesicle and the valve in it constituted a second mechanism operating on the input to the water clock to insure flow at a constant rate.

Negative feedback is a widely generalizable design principle, yet it had to be reinvented numerous times in human history (Mayr 1970). One of the more interesting reinventions was in James Watt's design of the centrifugal govern for the steam engine. The steam engine captures energy from combustion in the form of steam pressure and constrains the flow of steam to drive a flywheel to which appliances (e.g., sewing machines) are attached. One of the constraints within this controlled system is a valve that gates the flow of steam. In most applications, it is important that the power supplied to various appliances remains constant even as individual appliances make changing demands for power. This requires opening or closing the steam valve as needed. To register the engine's speed, Watt attached a spindle to the flywheel and attached arms to the spindle that would extend or retract based on centrifugal force. The angle of the arms represents (carries information) about the engine speed (Bechtel 2011), albeit in a non-intuitive way (Nielsen 2010). A linkage system connects

these arms to the steam valve in such a manner that the valve would gradually close as the arms raised (carrying information that the engine was running faster) and gradually open as the arms dropped (carrying information that the engine speed was running slower). The governor is a supplement to the basic controlled system of the steam engine. Via the gearing, the angle arms exercise a causal effect on the opening of the steam valve through which the steam flows, thereby operating on a constraint in the basic controlled system.

Watt's governor was such a success that it led James Clerk Maxwell (1868) to offer a mathematical analysis of governors. In the early twentieth century, negative feedback was employed in numerous designs of machines and became the foundation of the notion of circular causality celebrated by the cyberneticists (Wiener 1948). More recently, control theorists have developed more elaborate controllers that rely, for example, on forward models of the controlled system to determine alterations to the controlled system (Grush 2004). In all cases, the control mechanism operates on a constraint within the controlled system, altering the way in which energy is deployed to perform work.

4. Organisms Need Control Mechanisms to Maintain Themselves

Like human-made machines, mechanisms in organisms perform work by employing constraints to direct the flow of free energy. Most of these mechanisms perform work that is required to maintain the organism itself—build and repair its parts, capture energy in the form of ATP and make it available to other mechanisms, etc. Directing work to its own maintenance is required due to the fact that organisms are highly organized systems that, given the laws of thermodynamics, will inevitably break down (that is, they are dissipative structures). As we are all too aware, human-made machines break over time and have to be repaired (or, in our throwaway society, replaced). The need for repair is even greater in the case of biological organisms since they are largely soft systems, based on chemical processes, not hard or physical structures like most human-made machines. As a result, the forces holding the parts together are much weaker than in machines made out of materials such as wood or metal and these bonds are prone to break. On their own, the mechanisms within living organisms will degrade over time.

This need for repair has led theorists such as Robert Rosen (1985) to characterize living systems as self-repairing systems that he called *metabolism-repair* or *(M, R) systems*. The materials used to repair the system, as well as the energy needed to carry out the repair, can be recruited from outside the organism. Drawing on Aristotle's distinction between efficient and material causes, Rosen (1991) treats the repair system as materially open. But he argues that a (M, R) system must be closed to efficient causation—the

initiation of acts of repair must come from within the organism.[3] This means not only that the entity that initiates the repair of one item must be another component of the system, but also that the efficient cause for repairing that component when required must also come from within. The only way this is possible is if the sequence of efficient causes cycles back onto itself, resulting in a closed cycle. (This cycle doesn't present any challenge of backwards causation—the efficient cause of future repairs is the product of past actions within the organism.)

Not only must biological organisms repair their own mechanisms, they must also construct themselves to begin with. Every living organism starts as a cell that is produced by a division of an existing cell. Since cell division reduces the content of the cell, daughter cells must (re-)construct themselves. This led Maturana and Varela (1980) to characterize living organisms as self-constructing or *autopoietic* systems.[4] Proteins are the major constituents of living organisms. Accordingly, the mechanism of protein synthesis is one of the most important in the cell. The mechanism is usually viewed as stringing together amino acids into a polypeptide chair specified by the nucleic acid sequence in DNA. These polypeptide chains then fold into the required three-dimensional structure (often assisted by other proteins functioning as chaperones). This account brings out that one of the major constraints involved in the process of autopoiesis is the DNA-sequence the organism has inherited. (Accordingly, complex machinery is dedicated to checking and repairing DNA.) Focusing too much on DNA, however, can lead on to neglect the fact that free energy is required to perform the work of synthesizing proteins. Recognizing the need for free energy points us to a richer account of the constraints involved in living organisms. In particular, the enzymes that catalyze the steps from opening up the DNA to binding amino acids to one another, function as constraints directing energy in the manner needed to build proteins according to the constraint imposed by the DNA sequence.

Just as Rosen argued that the (M, R) system must be closed in terms of efficient causes, Pattee (see his papers collected in Pattee and Rączaszek-Leonardi 2012) and Alvaro Moreno and Matteo Mossio (2015) argue that the set of constraints that enable the organism to maintain itself must be closed in the sense that each constraint (enzyme, microtubule, etc.) must be constructed by mechanisms that rely on other constraints.[5] This requires a cycle of constraints: some of the (perhaps very indirect) products of one set of constraints are involved in construction of those constraints (on subsequent occasions). Accordingly, while a DNA sequence is one of the constraints involved in the synthesis of proteins, some of the proteins that are synthesized figure in the replication of DNA, detecting errors in replication, and carrying out repair. And among the polypeptide chains constructed in the ribosome are those that constitute the ribosome itself.

So far, I have focused just on the basic work that is required to make and repair biological organisms and not discussed control. If organisms existed

in an environment in which all the resources they need to perform the work to build and repair themselves were immediately available and each of the reactions occurred just when and where it was needed (just the right amount of ATP would be synthesized as needed to provide the free energy for protein synthesis, which in turn was perfectly coupled to the rates at which proteins are broken down), perhaps control would not be needed. Living systems, however, do not operate at anything like this level of accuracy. Errors that arise in processes such as transcription, translation, and folding of proteins must be corrected. Moreover, with perhaps the exception of single-celled organisms living in sulfur vents in the ocean, organisms must cope with varying environments and, in order to maintain themselves, have to adjust the activities they perform to the circumstances.

5. Maintaining Constancy of the Internal Environment via Control Mechanisms

The ability of organisms to maintain themselves has, at times, led biologists to reject the quest for mechanistic explanations. The vitalist Xavier Bichat (1805) opposed mechanistic explanations of biological phenomena because organisms (1) do not always behave in the same manner and (2) maintain themselves in the face of physical processes that would seem capable of destroying them (he characterized living systems as resisting death). Claude Bernard (1865) was one mechanist who took Bichat's contentions seriously and offered a framework for developing a mechanist answer. To account for the fact that organisms do not always respond to stimuli in the same way, he argued one must view the various mechanisms that constitute the organism as operating in what he termed the *internal environment*. This is the environment within the organism. Variation there would account for varied responses to external stimuli. To explain the resistance to death, he proposed that each mechanism is so designed to restore the constancy of the internal environment. Bernard, however, offered little insight into how each mechanism could operate to restore the constancy of the internal environment. Recognizing negative feedback as a design principle that enabled restoring a condition to its target state, Cannon (1929) offered several examples of how the autonomic nervous system employs negative feedback to maintain what he referred to as *homeostasis*. In the rest of this section, I describe two biological mechanisms in which feedback serves to maintain homeostasis, both serving to maintain an internal supply of ATP, the source of energy utilized in intracellular work.

The nineteenth century witnessed intense debates as to whether fermentation, the process of metabolizing glucose to yield alcohol and carbon dioxide, could be explained in terms of chemical reactions or required a whole living organism. This debate was largely resolved when Eduard Buchner (1897) observed the formation of carbon dioxide when he added

glucose to a cell-free extract and recognized this as a sign that fermentation was occurring without living cells. Although Buchner attributed this reaction to a single enzyme he named *zymase*, other researchers began to seek chemical intermediates, especially three-carbon compounds. Beyond the identification of pyruvate, the search for intermediates was largely foiled by the fact that most of the actual intermediates are phosphorylated compounds. Harden and Young's (1906) demonstration of the need to supply inorganic phosphate to sustain Buchner's reaction was puzzling since phosphates did not seem to appear in the products. Researchers soon recognized that fermentation was a variation on glycolysis, which figures in muscle contraction. Lundsgaard's (1930) discovery that phosphocreatine was the immediate source of energy for muscle contraction and Lohmann's (1929) discovery that the energy released in the oxidation of glucose was captured and stored for cell use in the phosphate bonds of adenosine triphosphate (ATP) revealed the importance of phosphorylated compounds at the end of glycolysis. Soon after, researchers showed that the intermediates in glycolysis were themselves phosphorylated and identified them. Since then, glycolysis has been viewed as a sequence of reactions as shown vertically in the center of Figure 12.2 (Bechtel 2006). Researchers recognized points at which ATP or ADP linked to the pathway (as source or recipient of phosphate bonds), but these were viewed as side processes off the main pathway.

Often glycolysis is presented as uncontrolled: as long as glucose is available, glycolysis proceeds. In fact, however, phosphorylated compounds, especially ATP, perform important regulatory roles, as shown by the reactions indicted by dashed lines on the right in Figure 12.2. Consider the third reaction in the pathway, which adds a phosphate group to fructose-6-phosphate to yield fructose-1,6-diphosphate. While ATP is an essential metabolite in the reaction itself, as it supplies the phosphate group, it is also an inhibitor of the enzyme. The enzyme phosphofructokinase-1 is an allosteric enzyme. Its conformation changes depending on whether it is bound to AMP or APD or to ATP. When bound to AMP or ADP, it phosphorylates fructose-6-phosphate more rapidly, at the expense of breaking down ATP to yield more ADP. This generates positive feedback. ATP, however, has the opposite effect, slowing the reaction. The physiological value of this design can be easily recognized. If the cell already has an ample supply of ATP, it would be wasteful to oxidize more glucose. It would be more efficient to maintain glucose in that form or convert it to glycogen until more ATP was needed.

As with negative feedback in human-made machines, negative feedback in glycolysis involves a secondary mechanism operating on the primary mechanism—the reaction pathway from glucose to lactate or alcohol. The control system is operating on the constraints (allosteric enzymes) of the main pathway, altering their operation. The next example is a little more complex since it is designed to register a condition in an organism's environment that is necessary before a mechanism can produce its desired effects.

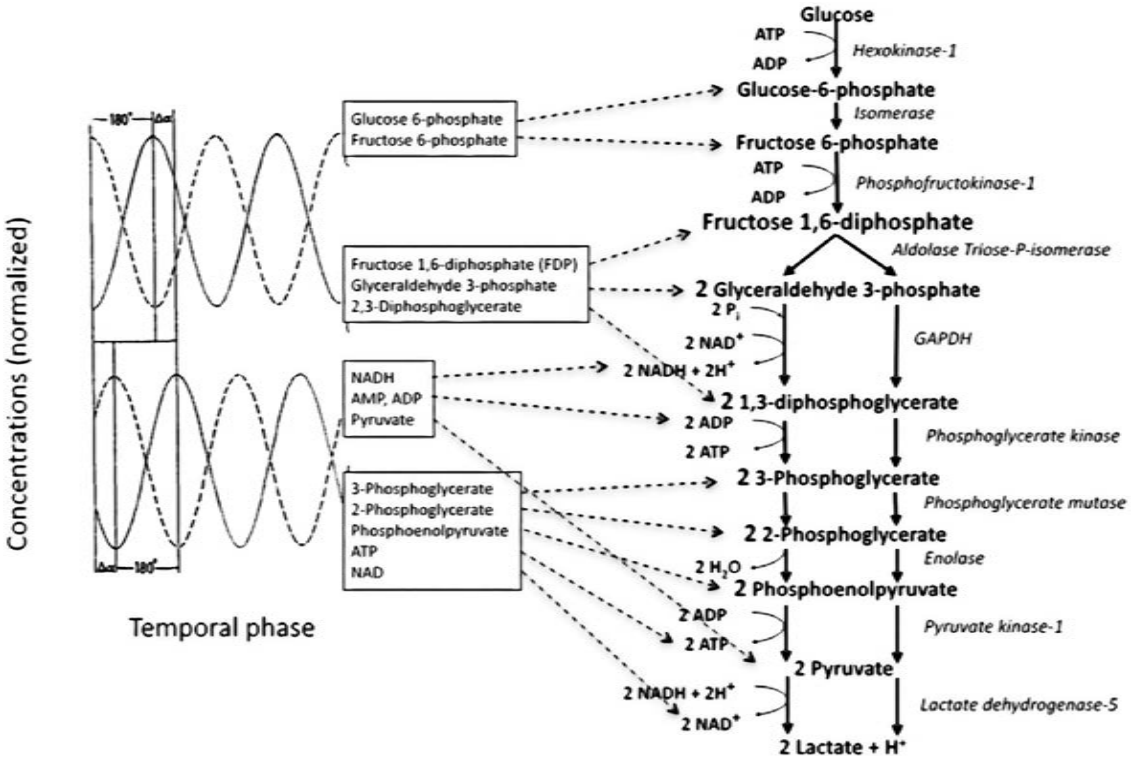

Figure 12.2 The glycolytic pathway is shown in the center with metabolites designated in dark text and enzymes in italics. Loops show where P_i, ADP, ATP, NAD^+, NADH and H_2O enter or leave the pathway. Dashed arrows and edge-ended lines on the right show feedback effects on enzymes (constraints) in the pathway. On the left are idealized graphs of the oscillation of the various intermediates, with dotted arrows linking the identification of the oscillating intermediates to where they appear in the pathway.

As in the case of glycolysis, it involves a control mechanism that operates on a constraint within the system that is being regulated.

In the 1930s, biochemical geneticists working with the bacterium *E. coli* discovered that the concentration of enzymes required for the metabolism of sugars such as galactose were not constant but would increase dramatically over time when the preferred sugar, glucose, was not available but galactose was. This process was originally designated *enzyme adaptation* and was thought to result in a modification of a precursor of the enzyme galactosidase when galactose was available. Monod, however, established that increased enzyme activity resulted from *de novo* synthesis of the enzyme from DNA. That is, it was by altering gene expression that control over the mechanism metabolizing galactose was achieved. This set Jacob and Monod (1961) on the quest that resulted in the discovery of one of the best-known control mechanisms in biology—the *lac operon*. The lac operon regulates the expression of three enzymes required to metabolize lactose, *lacZ*, *lacY*, and *lacA*. The key component in the operon is an allosteric enzyme, the *lac* repressor, which is constitutively produced by another gene, *lacI*. In the default state, it binds to the operator lying just in front for the three genes and largely blocks the RNA polymerase from initiating their transcription. The mechanism allows only a small, residual synthesis of *lacZ*. When lactose is present, the residual *lacZ* catalyzes the reaction producing allolactose from lactose. Allolactose binds to the *lac* repressor, altering its conformation so that it can no longer bind to the operator. This then allows the RNA polymerase to accelerate transcription of the three *lac* genes. An additional control mechanism prevents lactose from entering the cell whenever glucose is present, preventing this mechanism from accelerating the transcription of the *lac* genes except when lactose metabolism would be beneficial.

In this section, I have described two biological examples in which control mechanisms function to regulate the function of biological mechanisms so that they perform as needed to maintain the overall biological system. The glycolytic example involved negative feedback, in which ATP served to inhibit an operation in which it also functions as an input, thereby keeping ATP at constant levels in a cell. The *lac* operon uses feedback to detect the presence of lactose and accelerate the synthesis of the relevant genes when glucose is not available. In both cases the control mechanism operates on the constraints of another mechanism, adjusting its behavior so as to produce the results needed to maintain the constancy of the internal environment of the cell.

6. Using Oscillations from Negative Feedback to Control Timing of Operations

Engineers have long recognized that negative feedback often does not restore a system to its target value, but results in an oscillation around it. This is observed when a thermostat controls a furnace or air conditioner—first the

temperature exceeds the target, then it drops below the target, etc. Rather than stabilizing at the target temperature, it oscillates around it. In some cases oscillations generated by feedback mechanisms do dampen, but in other cases they sustain themselves. Negative feedback systems in biology also generate oscillations. Rather than just being a nuisance, as they often are in human-designed machines, oscillations are often employed as control systems in living organisms. Oscillations generate a repeating pattern of activity through time. The different activity states at different phases in the oscillation can be used to orchestrate operations of other mechanisms in time.

The glycolytic mechanism described above offers an example of feedback that generates oscillation. When Ghosh and Chance (1964) measured the concentration of NADH in their experimental preparation of yeast, they discovered it oscillated with a period of approximately one minute. Subsequently, Hess, Boiteux and Krüger (1969) demonstrated periodic oscillations in the concentrations of other reactants, with those generated in adjacent reactions generally being in phase with each other, but with phase shifts occurring at the phosphorylation of fructose-6-phosphate to fructose-1,6-diphosphate and the dephosphorylation of phosphoenolpyruvate to pyruvate (left side of Figure 12.2). They also observed a small phase delay between glyceraldehye-3-phosphate and 1,3-diphosphoglycerate, which is the step at which the oxidation reaction occurs. This phenomenon, known as *glycolytic oscillation*, is explained by the feedback loop involving the allosteric enzyme phosphofructokinase-1 discussed above. When AMP or ADP activates it, more 1,3-diphosphoglycerate is produced, which provides the input to subsequent reactions. Eventually NADH and ATP levels increase. The increased concentration of ATP serves to inhibit the reaction (and the declining concentration of ADP as it is phosphorylated to ATP also reduces its activating effect). As NADH is reduced in the formation of lactate and as ATP is consumed in performing different cell activities, the concentrations of NADH and ATP decline again.

Although glycolytic oscillation is readily demonstrated in laboratory conditions, it is uncertain whether oscillations occur under physiological conditions and whether it has any physiological functions (Richard et al. 1994; Richard et al. 1996). But there are many other negative feedback systems that produce oscillations in biological systems that have been demonstrated to perform regulatory roles. Among the best known are circadian oscillations, which are exhibited in a host of our own activities from sleep to athletic performance and in physiological processes such as metabolism and immune responses. Although in the following section I will identify an important role for neurons in circadian rhythms in animals, these rhythms are in fact generated within nearly all cells of our bodies. The core mechanism involves a transcription-translation feedback loop whereby the proteins PERIOD (PER) and CRYPTOCHROME (CRY) feed back to inhibit their own transcription. The steps in the process (accumulation of PER and

CRY in the cytoplasm, transport to the nucleus, binding to the proteins that activate transcription and removing them from the promoter, and then degrading) together take about 24 hours. The result is that concentrations of these (and several other proteins that are centrally involved in the mechanism) oscillate with a period of approximately 24 hours. Some of these oscillating proteins in turn serve as activators or inhibitors to other genes, causing them to be synthesized at appropriate times of day (e.g., proteins required for immune responses are synthesized at those times of day when we are most likely to encounter other people).

Although the core of the circadian clock mechanism involves negative feedback, it is a much more elaborate mechanism than the simple feedback loop in glycolysis. It involves a set of proteins (including many more than those indicated above) dedicated to the task of generating an oscillation with a period of about 24 hours (Reppert and Weaver 2002; Zhang and Kay 2010). Moreover, it is paradigmatically a control mechanism. It regulates a host of other mechanisms by sending signals that alter constraints (enzymes) within them. The various mechanisms that the circadian clock regulates can continue to function without it. Under such circumstances, these mechanisms cease to be coordinated with the light-dark cycle of our planet. This can have untoward effects on the health of the organism. The circadian system is at a higher level than these individual mechanisms and, when functioning properly, imposes top-down control that enables these mechanisms to generate their respective phenomena when appropriate for the organism.

7. Using Neurons to Realize Control Hierarchies

So far, I have illustrated the idea of top-down control without invoking the nervous system. This is appropriate since there is a great deal of control in single-cell organisms. Such control is required to integrate the activities of multiple mechanisms so that each performs as needed to maintain the existence of the organism. As we have seen, this control is typically exerted by altering the constraints in the mechanisms that channel and distribute energy into the performance of work. Moreover, yet higher levels of control can modulate lower levels. Within bacterial colonies there is differentiation of tasks between individual bacteria, and signaling systems exist that enable the colony to alter the operation of control mechanisms in individual bacteria. This differentiation of function and accompanying hierarchical control becomes even more manifest in multi-cell organisms. One of the central modes of control is achieved through the emergence of specialized cells, neurons, with long projections (axons and dendrites) from the cell body. Neurons conduct electrical charges along these projections until a synapse is reached. There they release transmitters that can excite other cells. (In some cases, electrical signals are directly communicated to other cells through what are known as *gap junctions*.)

Much thinking about neural control systems has adopted Charles Scott Sherrington's (1923) view of the nervous systems as largely a reactive system in which sensory stimulation initiates a sequence of neural activity culminating in a motor response. On this view, the importance of the nervous system is to enable organisms to respond appropriately to conditions in their internal or external environments. In simple reflex cases, nerves from sensors control motor outputs, enabling them to respond appropriately to stimulus condition. When determining the needed response is more difficult, a network of neurons intervenes. If appropriately configured, such networks can learn to respond differentially to the encountered circumstances. To then exercise control, some of the neurons must connect directly to other tissues that perform physiological processes or motor actions. On this view, which Keijzer (2015) characterizes this as an input-output conception of the nervous system, brains are hierarchies of complex networks. Networks higher in the hierarchy control those lower and the network at the lowest level controls the motor outputs.

On this reactive input-output view of the nervous system, one would expect an organism to remain passive until it received input. But observing any animal confounds this assumption—animals are endogenously active. This is true not just of animals; even single-celled organisms are characteristically active both in carrying out basic life functions and in moving through space. Reversing the usual perspective, activity might be viewed as the default state with special arrangements required in order to stop activity. From this perspective, what the nervous system must do is constrain endogenous activity so as to enable coordinated action. (Keijzer thus contrasts the input-output view with what he terms the *coordination view*. For him, the first neurons to evolve served to coordinate contractile tissues so as to generate locomotion. Even if coordinating motility was the original role of neurons, they provided as well a basis for coordination of other activities, including more basic physiological functions.)

Fundamental to the coordination view is the contention that the systems that need coordination are endogenously active. A similar assumption is appropriate for the neurons that specialize in coordination. Within Sherrington's laboratory, Thomas Graham Brown (1914) offered just such a view of the nervous system. Although ostensibly investigating reflexes, he began to attend to the endogenous rhythmic activity that persisted even in deafferented legs in rabbits and other mammals. This research received little uptake at the time. It was revived, however, in research on central pattern generators—networks of neurons that are active in generating cycles of motor activity without external stimulation (Wilson (1961)). More recently, central pattern generators have been found to control a great variety of other neural activity including visual and olfactory processing and cognitive activities including memory formation.

Neural pattern generators require ongoing physiological activity within neurons (resulting from constraining the release of free energy within them)

and a mode of organization (itself either within or between neurons) through which the products of these activities constrain others activities. Neurons and the nervous system are endogenously active systems (Bechtel 2013) that can then control other mechanisms. The mechanism for generating circadian rhythms discussed above is one example of an endogenously active control system, but there are many others found in the nervous system. For these endogenously active neural mechanisms to control other biological mechanisms, they must affect constraints in these mechanisms. Sometimes a complex set of operations intervenes between the neural controller and the controlled organs. In the case of muscles, for example, those neurons whose axons synapse onto muscles release neurotransmitters that bind to receptors on the muscle. This generates an electrical current within the muscle cell that leads to a release of calcium from the sarcoplasmic reticulum into the cytoplasm. There the calcium reacts with troponin, causing it to bind to tropomyosin, which was blocking the binding sites between actin and myosin. This then permits the cycling of cross-bridges that cause actin and myosin filaments to pull each other in. This continues until the electrical current ceases, stopping the release of calcium. In this scenario, different constraints are modified in sequence resulting in releasing the endogenous interaction of actin and myosin filaments.

Once neurons evolved as cells that could control the operation of other cells by altering constraints in them, the path was open for creating a hierarchy of such constraints. Constraints in individual neurons could be modified by activity in networks of neurons, and yet higher-level networks could operate on neurons in these networks. I will illustrate this potential by returning to the example of the circadian feedback mechanism operative in individual cells, including individual neurons. In animals, either collections of neurons (e.g., in fruit flies) or whole nuclei (in mammals) assume a regulatory role with respect to the oscillators in individual cells. In mammals, a structure known as the suprachiasmatic nucleus (SCN) performs this function. If the SCN is surgically removed, the animal ceases to exhibit circadian rhythms in behavior or in physiological function (Moore and Eichler 1972). If slices from the removed SCN are maintained in an appropriate medium, the neurons continue to generate circadian rhythms (Herzog et al. 2004), indicating that slices of the SCN can function autonomously. If, however, SCN neurons are dispersed so that many of the connections between them are lost, individual cells still oscillate, but with substantially varying periods, ranging from 21.25 to 26.25 hours with a SD of 1.2 hours (Welsh et al. 1995). Since individual oscillations are out of phase with each other, there is no detectable rhythm in the overall populations. Given that regular rhythms are found in normal SCN tissue in which cells communicate, the communication must synchronize the endogenous oscillations. Thus, collectively the cells of the SCN regulate each other's behavior, resulting in far more reliable timekeeping than individual neurons can produce. This top-down effect from the population to the individual results from many individual

SCN neurons sending signals to which others can respond by advancing or delaying their own oscillation.

As I noted above, individual cells in mammals possess the requisite mechanism for generating circadian rhythms. What they lack is the ability to synchronize the rhythms in individual cells. This requires the SCN, which functions as a controller on their rhythms. How a signal is communicated from the SCN to other cells of the organism is not understood. When Ralph et al. (1990) removed the native SCN in a hamster and inserted the SCN from a mutant strain that exhibited short periods into a ventricle, they succeeded in restoring some circadian behavior but with a short period. Since the inserted SCN did not make neural projections, its effects on other tissues must have been through hormones. But the fact that not all behavioral or physiological rhythms could be restored suggests that the effect of the SCN on other mechanisms may require neural transmission.

Since circadian rhythms, as the name implies, have a period of only approximately 24 hours, it is important that SCN cells also be entrained to the external environment by sensory information. Otherwise, after a few days an organism will be out of phase with the light-dark cycle in its environment. In fact, one of the initial clues that the SCN was the central clock was that it receives projections from the retina. After the details of the circadian mechanism were discovered, researchers identified the pathway by which the signal from the retina serves to enhance the concentration of PER within a population of SCN cells. If the signal is received around expected dawn, when PER levels are beginning to increase, the signal serves to advance the phase of the oscillation. If, on the other hand, it is received around expected dusk, it serves to delay the phase. The retina thus provides higher-level control over the SCN, which in turn regulates individual cells throughout the body that directly affect the transcription of many proteins which figure in basic activities of organisms. Moreover, one can even view the retina as part of a higher-level control circuit that includes the locomotor system and decision-making operations since exposure to light is also affected by the behavior of the organism. This is particularly true of nocturnal organisms, which must exit their burrows to receive light input. Such higher-level intervention is also a factor in us: when humans expose themselves to light at night (e.g., in performing shift work), they cause their circadian rhythms to be desynchronized from the light cycle in their environment. This in turn frequently results in obesity, diabetes and various cancers.

The circadian system is just one example of a hierarchical control system realized through neurons. There is not space to describe others in detail, but the basic pattern is the same. As research with decorticated animals makes clear, basic motor activity is retained, but less coordinated, when neural control is removed. Sub-cortical brain regions provide a great deal of the needed control. Cortex serves as a higher-order control system that is linked to subcortical ones through numerous loops involving projections both up to cortex and back down to sub-cortical areas. At each level, researchers are

identifying complex mechanisms that maintain their own dynamical behavior while modulating constraints in ones lower in the constraint hierarchy.

8. Conclusion: Hierarchical Control as Top-Down Causation

My aim in this paper has been to articulate a notion of top-down causation appropriate to control hierarchies in biology. I have characterized mechanisms as collections of parts that through their operations constrain the flow of free energy so as to perform work. Biological systems are often viewed as networks. Modules (clusters of units) in these networks often correspond to mechanisms as more traditionally characterized through research that first identifies mechanisms and decomposes them (Bechtel and Richardson 1993/2010; Craver and Darden 2013). Clustering provides a means of differentiating mechanisms from the rest of the components of the network. The interactivity within mechanisms often yields complex dynamical activity. Connections to nodes outside the mechanism then play critical roles in determining the behavior of modules. Some of these connections simply involve the transfer of matter or energy between mechanisms, providing the resources each mechanism needs to perform work. But others serve to control activity within them. In many cases, this control is exercised by higher-level, dedicated control mechanisms.

Control of mechanisms is extremely important in living organisms since they must both construct and maintain themselves as organized systems consisting of multiple mechanisms that are subject to degradation. Mechanisms must perform the activities of constructing and repairing themselves in varying environments that place different demands on the mechanisms that constitute them. Control of a mechanism is achieved by changing the constraints that direct energy to perform work. Some constraints in mechanisms are fixed, but others can be modified. The latter provide the opportunity for control—altering these constraints causes the mechanism to generate different behavior. Neurons are not necessary for exercising control—mechanisms within single-celled organisms are controlled through chemical signals. These enable the organism to maintain itself despite highly varying conditions in which it must function. But neurons provide a potent way to exercise coordinated control over a variety of different mechanisms. As well, they afford the development of a hierarchy of control systems enabling greater ranges of control, including the sorts of control humans can realize in their voluntary actions (Bechtel 2008).

Control systems are appropriately viewed as at a higher level than the systems they control. They operate on a controlled system (mechanism) by altering parts that serve as constraints within it and thereby alter its behavior. Control systems are distinct from other mechanisms that have causal effects on the controlled mechanism in that they do not supply the matter or energy needed for the controlled mechanism to perform work. Both the controlled

and the control mechanisms require appropriate material inputs and free energy. Ultimately, these are extracted from the environment through specialized mechanisms. Every mechanism is characterized by constraints that enable work to be done using the free energy that is available to it. The work that a control mechanism does is to alter constraints within the mechanism being controlled, thereby affecting how the controlled mechanism uses its matter and energy. In human-made machines, controllers serve to keep the machines functioning as their designers intended. In the case of biological mechanisms, control mechanisms keep other mechanisms operating in the manner need for the organism to construct its own mechanisms and maintain them as they degrade. Like all mechanisms, control mechanisms are made through the operation of other mechanisms within the organism or its parent. They are differentiated as at a higher level because they operate on the constraints in lower-level mechanisms. This notion of top-down causation is both principled and needed to understand living organisms.[6]

Notes

1 Although I am developing the concept in a slightly different way, my discussion of control hierarchies is inspired by Howard Hunt Pattee (1970), (1972). I am not arguing that hierarchical control is the only useful notion of top-down causation that can be applied in neuroscience, but only that it is a very important notion.

2 Gibbs free energy is the thermodynamic potential that specifies how much work can be performed by a system at constant temperature and pressure. Formally: Gibbs free energy = enthalpy—temperature (Kelvin) x entropy.

3 If the series of repair operations were not closed to efficient causation, the result would be an infinite regress—each repair operation would be dependent on another to repair it. Successful repair would rely on just the right string of causal processes and could not be relied upon. In the case of human-made machines, the repairperson resides outside the mechanism, but if the repair process is to be reliable, the repair system (repair people, parts supply, etc.) must be maintained as a self-sustaining system (that trains new repair people, orders the right parts, etc.). Biological organisms typically don't have an external repair system that they can be rely on. Sometimes a symbiotic organism (e.g., a bacterium residing in a multi-celled organism) may perform repair activities; in this case the repair system extends outside the organism, but it must still be closed (e.g., in the coordination between host and bacterium) if the host is to be able to rely on the bacterium for repair.

4 In constructing themselves, organisms rely on material and free energy from outside. But these must be utilized appropriately to create a new organism and the mechanisms directing this use must reside within the living organisms themselves (although they may rely critically on appropriate conditions in their environment to carry out the needed operations). Gánti (1975) proposed a mechanism he called a *chemoton* as the simplest chemical system able to construct itself. It consisted of a membrane that controlled access to the internal environment, a metabolic system to transform inputs, and a regulatory system that determined what metabolic operations to perform.

5 Like Rosen and Varela and Maturana, Moreno and Mossio argue that closure of constraints is required if an organism is to reliably maintain itself and not dissipate. In some cases organisms can off-load this responsibility if the environment

can be counted upon to provide the necessary constraint. For example, we have off-loaded the synthesis of some essential molecules—vitamins—to other organisms that provide our foodstuffs. But as vitamin deficiency diseases make clear, such reliance can prove fatal when the source of food changes. Humans have addressed this by creating an industry to provide vitamins, but this is just an extension of what must remain a closed system of constraints to include constraints outside the organism.

6 I thank Jason Winning for many productive discussions on topics related to this paper and very helpful comments on a previous draft. I also thank Alvaro Moreno and his collaborators (especially Kepa Ruiz-Mirazo and Leonardo Bich) for discussions that helped inspire the views presented here.

References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., Sherlock, G. (2000). "Gene Ontology: Tool for the Unification of Biology". *Nature Genetics*, 25: 25–29

Bandyopadhyay, S., Kelley, R. M., Krogan, N. J., Ideker, T. (2008). "Functional Maps of Protein Complexes from Quantitative Genetic Interaction Data". *Plos Computational Biology*, 4

Bandyopadhyay, S., Mehta, M., Kuo, D., Sung, M. K., Chuang, R., Jaehnig, E. J., Bodenmiller, B., Licon, K., Copeland, W., Shales, M., Fiedler, D., Dutkowski, J., Guenole, A., van Attikum, H., Shokat, K. M., Kolodner, R. D., Huh, W. K., Aebersold, R., Keogh, M. C., Krogan, N. J., Ideker, T. (2010). "Rewiring of Genetic Networks in Response to DNA Damage". *Science*, 330: 1385–1389

Barabási, A. L., Bonabeau, E. (2003). "Scale-Free Networks". *Scientific American*, 288: 50–59

Bechtel, W. (2006). *Discovering Cell Mechanisms: The Creation of Modern Cell Biology*. Cambridge: Cambridge University Press

——— (2008). *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. London: Routledge

——— (2011). "Representing Time of Day in Circadian Clocks". In: Newen, A., Bartels, A., Jung, E. M. (eds.), *Knowledge and Representation*. Palo Alto (CA): CSLI Publications: 129–162

——— (2013). "The Endogenously Active Brain: The Need for an Alternative Cognitive Architecture". *Philosophia Scientiae*, 17: 3–30

——— (forthcoming). "Explicating Top-Down Causation Using Networks and Dynamics". *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., Richardson, R. C. (1993/2010). *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge (MA): MIT Press. 1993 edition published by Princeton University Press

Bernard, C. (1865). *An Introduction to the Study of Experimental Medicine*. New York: Dover Publications

Bichat, X. (1805). *Recherches Physiologiques sur la Vie et la Mort*. 3rd Edition. Paris: Machant

Brown, T. G. (1914). "On the Nature of the Fundamental Activity of the Nervous Centres; Together with an Analysis of the Conditioning of Rhythmic Activity in Progression, and a Theory of the Evolution of Function in the Nervous System". *The Journal of Physiology*, 48: 18–46

Buchner, E. (1897). "Alkoholische Gärung ohne Hefezellen (Vorläufige Mittheilung)". *Berichte der deutschen chemischen Gesellschaft*, 30: 117–124
Cannon, W. B. (1929). "Organization of Physiological Homeostasis". *Physiological Reviews*, 9: 399–431
Craver, C. F., Bechtel, W. (2007). "Top-Down Causation without Top-Down Causes". *Biology and Philosophy*, 22: 547–563
Craver, C. F., Darden, L. (2013). *In Search of Mechanisms: Discoveries across the Life Sciences*. Chicago: University of Chicago Press
Gánti, T. (1975). "Organization of Chemical Reactions into Dividing and Metabolizing Units: The Chemotons". *BioSystems*, 7: 15–21
Gene Ontology Consortium (2015). "Gene Ontology Consortium: Going Forward". *Nucleic Acids Research*, 43: D1049–D1056
Ghosh, A. K., Chance, B. (1964). "Oscillations of Glycolytic Intermediates in Yeast Cells". *Biochemical and Biophysical Research Communications*, 16: 174–181
Grush, R. (2004). "The Emulation Theory of Representation: Motor Control, Imagery, and Perception". *Behavioral and Brain Sciences*, 27: 377–396
Harden, A., Young, W. J. (1906). "The Alcoholic Ferment of Yeast-Juice, Part II: The Coferment of Yeast-Juice". *Proceedings of the Royal Society, London*, B78: 369–375
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: 35–46
Hess, B., Boiteux, A., Krüger, J. (1969). "Cooperation of Glycolytic Enzymes". *Advances in Enzyme Regulation*, 7: 149–167
Jacob, F., Monod, J. (1961). "Genetic Regulatory Systems in the Synthesis of Proteins". *Journal of Molecular Biology*, 3: 318–356
Keijzer, F. (2015). "Moving and Sensing without Input and Output: Early Nervous Systems and the Origins of the Animal Sensorimotor Organization". *Biology & Philosophy*, 30: 311–331
Kim, J. (1998). *Mind in a Physical World*. Cambridge (MA): MIT Press
Lohmann, K. (1929). "Über die Pyrophosphatfraktion im Muskel". *Naturwissenschaften*, 17: 624–625
Lundsgaard, E. (1930). "Untersuchungen über Muskelkontraktionen ohne Milchäurebildung". *Biochemische Zeitschrift*, 217: 162–177
Machamer, P., Darden, L., Craver, C. F. (2000). "Thinking about Mechanisms". *Philosophy of Science*, 67: 1–25
Maturana, H. R., Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: Reidel
Maxwell, J. C. (1868). "On Governors". *Proceedings of the Royal Society of London*, 16: 270–283
Mayr, O. (1970). *The Origins of Feedback Control*. Cambridge (MA): MIT Press
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
Moreno, A., Mossio, M. (2015). *Biological Autonomy: A Philosophical and Theoretical Inquiry*. Dordrecht: Springer
Nielsen, K. (2010). "Representation and Dynamics". *Philosophical Psychology*, 23: 759–773
Pattee, H. H. (1970). "The Problem of Biological Hierarchy". In: Waddington, C. H. (ed.), *Towards a Theoretical Biology 3: Drafts*. Edinburgh: Edinburgh University Press: 117–136
——— (1972). "The Nature of Hierarchical Controls in Living Matter". In: Rosen, R. (ed.), *Foundations of Mathematical Biology, Vol. 1: Subcellular Systems*. New York: Academic Press: 1–22

Pattee, H. H., Rączaszek-Leonardi, J. (2012). *Laws, Language and Life: Howard Pattee's Classic Papers on the Physics of Symbols with Contemporary Commentary by Howard Pattee and Joanna Raczaszek-Leonardi*. Dordrecht: Springer

Ralph, M. R., Foster, R. G., Davis, F. C., Menaker, M. (1990). "Transplanted Suprachiasmatic Nucleus Determines Circadian Period". *Science*, 247: 975–978

Reppert, S. M., Weaver, D. R. (2002). "Coordination of Circadian Timing in Mammals". *Nature*, 418: 935–941

Richard, P., Diderich, J. A., Bakker, B. M., Teusink, B., van Dam, K., Westerhoff, H. V. (1994). "Yeast Cells with a Specific Cellular Make-Up and an Environment That Removes Acetaldehyde Are Prone to Sustained Glycolytic Oscillations". *FEBS Letters*, 341: 223–226

Richard, P., Teusink, B., Hemker, M. B., Dam, K. V., Westerhoff, H. V. (1996). "Sustained Oscillations in Free-Energy State and Hexose Phosphates in Yeast". *Yeast*, 12: 731–740

Rosen, R. (1985). "Organisms as Causal Systems Which Are Not Mechanisms: An Essay into the Nature of Complexity". In: Rosen, R. (ed.), *Theoretical Biology and Complexity: Three Essays on the Natural Philosophy of Complex Systems*. New York: Academic Press: 165–203

——— (1991). *Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*. New York: Columbia University Press

Sherrington, C. S. (1923). *The Integrative Action of the Nervous System*. New Haven (CT): Yale University Press

Watts, D., Strogratz, S. (1998). "Collective Dynamics of Small Worlds". *Nature*, 393: 440–442

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: 697–706

Wiener, N. (1948). *Cybernetics: Or, Control and Communication in the Animal and the Machine*. New York: Wiley

Wilson, D. M. (1961). "The Central Nervous Control of Flight in a Locust". *Journal of Experimental Biology*, 38: 471–490

Zhang, E. E., Kay, S. A. (2010). "Clocks Not Winding Down: Unravelling Circadian Networks". *Nature Reviews Molecular and Cell Biology*, 11: 764–776

13 Early Complexity in Human Development

François Jouen and Michèle Molina

Developmental sciences cannot avoid the question of the origins and nature of knowledge. Piaget clearly placed this issue at the heart of his genetic approach. Piaget (1936, 1937) strongly opposes the idea of predetermined knowledge to that of knowledge as actively constructed by the subject. Piaget clearly fixed the objectives of genetic epistemology: to account for the construction of non-preformed structures from which cognitive mechanisms would emerge. Well before the explosion of research on early skills during infancy, this constructivist line has been challenged by nativist approaches: just remember the confrontation between Piaget and Chomsky in Royaumont in 1975 on the question of language acquisition, during which Chomsky clearly confesses that his nativist conceptions are completely opposed to Piaget's constructivism. Since 1975, the question of the origins of cognition was almost exclusively limited to a debate between nativists and empiricists. Both approaches, radically antagonistic, nevertheless agree on two points: the representative basis of knowledge and a rejection of the fundamental assumption of the constructivist position, according to which cognition is deeply rooted in sensorimotor activity. A positive consequence of the nature-nurture debate was to greatly increase the number of studies on cognitive development in infants and children. Whatever theoretical model they adopted, researches concerning younger and younger infants started with the main goal of getting closer to the early development of knowledge. In this perspective, birth was taken as the zero state of the initial development of cognition. The implicit idea was that dating a cognitive skill closer to birth could give access to what is *biologically determined*. The identification of early cognitive competence in the newborn infant was accompanied by animated debates about the *predetermined nature* of knowledge: just remember the vigorous discussions that followed the demonstration of neonatal imitation by Meltzoff and Moore (1977). The Chomskian concept of mental organs characterizes the innate cognitive framework of nativism: like all physical organs, mental organs are genetically determined and are species-specific. By referring to the well-known poverty of the

stimulus argument, nativists oppose the scarcity of the stimulus and of the perceptual functions that ensure its processing to the complexity of mental structures that are defined as intrinsic, idiosyncratic, rich and various. For empiricists, who are faithful to Aristotelian tradition, what is in the mind was previously in the senses.

Despite the theoretical interest arising from the opposition between nativist and dualistic approaches, it is clear that the debate is still in progress regarding the origin of knowledge in infants. Criticisms arising from renowned biologists such as François Jacob, who suspected Piaget of neural Neo-Lamarckism, renew the question of the origins of knowledge by referring to theoretical models based on embryogenesis and probabilistic epigenesis. Piaget (1967) was among the first researchers who argued that the universality of a behavior does not necessarily imply genetic transmission. He rather suggested that brain structures and associated mental functions can exhibit *self-stabilization* as a consequence of interactions between the genetic heritage of a species and individual experience. This idea is obvious in various biological models, such as the theory of selective stabilization of synapses proposed by Changeux, Courrège and Danchin (1973), the theory of developmental psychobiological systems proposed by Gottlieb (1991), or the theory of neuronal groups stabilization proposed by Edelman (1992). Curiously, these approaches were, with rare exceptions (Hadders-Algra 2000, 2002; Jouen and Molina 2007), very seldom applied to early cognitive development, which is puzzling, since they offer a promising alternative to the debate between nativist and empiricist approaches. The main objective of this article is to examine the contribution of these recent biological approaches to the question of the origins of knowledge in infancy.

1. The Idea of Early Competence

Initially developed in 1936, Piaget's theory has long prevailed as the exclusive model of children's cognitive development. This supremacy was sometimes vehemently challenged during the 1970s and the 1980s by many studies on perceptual and cognitive skills in young children. These studies have highlighted the idea of a competent infant endowed with cognitive skills observable earlier than previously assumed by Piaget's model. The idea of the competent infant is deeply related to the development of two techniques: the preferential-looking paradigm and the habituation paradigm. The former (Fantz 1956) consists in observing the distribution of visual fixation durations on two targets that are side by side, varying in one dimension (shape, color, size, arrangement of elements, etc.). If infants look longer at one target than at the other, researchers conclude that infants do not consider the two targets equivalent: infants have detected and discriminated the difference between the targets by coding the information

contained in each target. By abuse of language, researchers talk of visual preference. However, concluding that there is visual preference is possible only if infants look at each target at least once. If infants look at no target or if they look equally at the two targets, they do not show visual preference. Yet, in this case, experimenters cannot conclude that infants were not able to discriminate between the two targets. To avoid such difficulties, the technique of habituation is frequently used (Heering 2010). The habituation paradigm (when learning duration depends on infant's activity) or familiarization (when the duration of learning is determined by the investigator) consists in recording the reduction of visual fixation durations in relation to the repeated presentation of the same stimulus. After this training period, the familiar target is presented, in competition with or in alternative to a new target, during a test period. Generally, infants tend to look longer at the new target than at the familiar one. The *response to novelty*, observed during the test period, demonstrates that infants are able to compare the new stimulus to the stored familiar target, and to perfectly discriminate one from the other.

These techniques, initially used to investigate perceptual skills, have revealed that newborn infants have a significant number of perceptual skills that, though tenuous, demonstrate the existence of functional corticalization since birth. This conclusion sharply contrasts with the conclusions of authors who claimed, for instance, that the visual cortex was not functional during the first weeks of life (Bronson 1974; Johnson 1990). As shown by Slater, Morison and Somers (1988), a critical test of cortical function is the perception of orientations. Following habituation to a diagonal grating tilted at 135° or 45°, newborns look longer at a mirror-image grating. Similarly, works on face perception (Schonen, Mancini and Liegeois (1998)) attest the involvement of the cortex in the control of neonatal cognitive activity: the preference for the maternal face is the result of an extremely fast perceptual learning in contact with the face of the mother (Pascalis et al. 1995).

Demonstrating cortical functions in newborns necessarily means that, from birth, they access a complex visual world that cannot be reduced to a set of bright spots, each being present during fixation and then forgotten. Various studies have shown that newborns demonstrate shape constancy (i.e., the ability to recognize the shape of an object despite changes in its orientation) and size constancy (i.e., the ability to perceive the objective size of an object despite changes in its distance from us), as shown by the works of Slater and Morison (1985) and Slater et al. (1991). The existence of perceptual constancies allows newborns to access a stable world, coherent and composed of tangible units. Contrary to Piaget's postulates, the visual world of the newborn is not limited to a non-structured two-dimensional world that only manual exploration can make three-dimensional. From birth, the infant is able to discriminate

two-dimensional stimuli from three-dimensional stimuli. In a three-dimensional environment, the newborn is able to visually process stimulus compounds differing in orientation, size and color. The experiment conducted by Slater et al. (1991) gives an example of such competence. Newborns are familiarized with two alternately presented stimuli that differ in color and orientation (for example, a green vertical bar and a red oblique bar). During the test, they are presented with a familiar stimulus or a new stimulus created by recombining features of the stimuli used for familiarization (a red vertical bar and a green oblique bar). Newborns look longer at the new stimulus: this finding demonstrates that, during habituation, infants have processed the relation between shape and color. Newborns do not process separate components: they are able to combine different properties of objects, which is fundamental to ensure the visual perception of objects (Triesman 1986).

The study of these early perceptual skills has been supplemented with the description of early cognitive skills. These studies are merely based on the habituation paradigm. Although habituation technique relies on perceptual discrimination, some researchers have assumed that *conceptual* habituation is also conceivable. In the 1980s, habituation was used to probe conceptual knowledge in infancy, such as the knowledge of the permanence of objects (Kellman and Spelke 1983; Baillargeon, Spelke and Wasserman 1985; Baillargeon and Graber 1987) or of the concept of number (Wynn 1992). Very elegant research, in which only a conceptual dimension is modified between the habituation and the test phases, was conducted on infants of less than four months of age, i.e., infants who do not yet have eye-hand coordination and who are not yet able to grip and manipulate objects. The logic of these experiments is as follows. If infants look longer at the test items, this means that they are sensitive to the conceptual dimension manipulated by the researcher: They react to conceptual novelty and not only to perceptual novelty. However, infants also have the opportunity to respond to the perceptual changes that necessarily occur when manipulating the conceptual dimension during the test phase. Consequently, researchers oppose conceptual novelty to perceptual familiarity or perceptual novelty to conceptual familiarity.

Take, for example, Baillargeon's (1987) famous experiment of the drawbridge—which tests knowledge about the concept of object. How do infants know that the movement of physical objects is constrained by the principle of object solidity? In this research, four-month-old infants are habituated to a screen motion that rotates back and forth 180 degrees over repeated trials. At the end of the habituation period, a real and visible box is placed behind the screen. Initially, when the screen is flat against the table, infant can see the box. Once the screen starts rotating, it progressively prevents the infant from seeing the box. Two test events are then presented to babies. The *possible* test event presents the screen that rotates only 112 degrees and is locked by the box, which is not visible to

infants at this moment. The *impossible* test event shows the screen that does a complete 180-degree rotation despite the presence of the obstacle (which is no longer visible when the screen reaches 112 degrees of tilt). Infants assigned to a control group are habituated to the same sequence and are tested using the same test events presented to infants placed in the experimental condition, with only one difference: no obstacle is placed along the path of the screen. In this experiment, infants can react either to the perceptual change (i.e., they can detect that the screen no longer rotates 180 degrees but 112 degrees) or to the conceptual dimension manipulated by the researcher (i.e., they can be surprised by the rotation of the screen despite the presence of an obstacle). Results show that infants in the experimental condition look significantly longer at the impossible event than at the possible event. Infants in the control condition do not prefer any of the two tests events. The author concludes that, as soon as the age of four months, infants are able to separate perceptual aspects to respond to conceptual elements. Baillargeon (1987) considers that, from the age of four months, infants know that the obstacle continues to exist behind the screen, which clearly shows early knowledge of object permanence. Moreover, babies would develop cognitive activity allowing them to infer that the screen rotation must necessarily be blocked by the obstacle.

Data obtained from the numerous studies that have used the method of violation of expectancies have shown, sometimes quite spectacularly, that young infants of four months of age or younger are able to react to perceptual differences and, most importantly, to gain some knowledge about their environment in spite of their reduced sensorimotor activity. This conclusion is in deep opposition to the statement of Piaget's theory, according to which such knowledge can be built only from the training of sensorimotor activity. This research, sometimes conducted as soon as birth (for a review see Slater (1995), (1997); Slater and Johnson 1998), has undoubtedly revealed that, before they can physically act on their environment, infants possess sufficiently developed cognitive skills that can help them to make sense of their environment (Bryant and Trabasso 1971; Gelman 1969). This challenge to the foundations of Piaget's theory has directly led to the question of the *nature* of early skills authenticated by researchers. If the sensorimotor system has no role in cognitive development and if the subject does not build knowledge from the sensorimotor actions he performs in environment, this necessarily means that knowledge exists prior to action. As a consequence, newborn cognitive development was no longer considered in terms of the Piagetian statement, but in terms of the nativist position. Newborns were thus endowed with cognitive processes and knowledge innately determined. In this context, the cognitive competences revealed by this research could either be understood as reflecting initial knowledge or as revealing an infant's initial cognitive endowment.

2. From the Dissociation of Competence and Performance to the Earliness Argument

The rejection of Piaget's theory by nativists was also based on the concepts of competence and performance initially introduced by Chomsky (1965) in order to defend a nativist approach of language development. According to Chomsky, performance does not give direct access to competence, since many factors can potentially limit the expression of a skill. Chomsky defends a kind of nativism closer to a predeterministic approach based on the argument of *the poverty of the stimulus* originally developed by Plato. Arguments in support of the poverty of the stimulus are based on the idea that mental structures and organized complex knowledge cannot be built on sensory experiences that are unreliable, incomplete and sometimes false. Language experience is sometimes erroneous and incomplete, but extremely varied. However, a child produces many new phrases he has never heard before. This ability would demonstrate that the child does not learn grammar simply by repeating what he heard. Chomsky finds here the confirmation of the fundamental distinction between perceptual and cognitive processes: cognitive activity necessarily contains more information than perceptual inputs. The latter must be interpreted by predetermined rules. Based on this argument of the poverty of the stimulus, Chomsky claims that children's grammatical capacities are necessarily innate and exist in the form of a set of genetically encoded and biologically inherited rules. The existence of innate grammar rules explains why some aspects of language would develop with incredible consistency despite the originality of each child's experiences and environmental peculiarities. These innate rules also explain why children's language skills transcend performance and, conversely, why performances do not reflect skills.

The dissociation between competence and performance has resulted in the nativist research program centered on the identification of increasingly early skills, up to the point that they could not be explained by any other conventional developmental theory than the one proposed by the nativist approach. Fischer and Bidell (1991) named this research strategy the "*argument of precocity*", through which nativists conclude that, if infants exhibit cognitive skills that could not be learned through or built by sensorimotor activity, such skills are innate (or innately constrained). This persistent quest for precocity continues today through the development of research using the habituation paradigm in preterm infants (Lejeune et al. 2010). This research shows the preterm infants' ability to process the shape of objects in a similar way to what is observed in the newborn at term, as if the special experience of prematurity for the development of cortical functioning had no influence. The development of neuroimaging techniques such as infrared spectroscopy remains in the line of quest for precocity. However, although demonstrating changes in the local concentration of oxyhemoglobin and deoxyhemoglobin in the somatosensory cortex of

28-week premature babies in response to tactile stimulation is a central issue to investigate the maturation of the cerebral cortex (Roche Labarbe et al. 2014), this gives little information on how the premature newborn gains access to a world of objects.

The argument of precocity is underpinned by a fundamental premise defended by the nativists: experience is not sufficient to account for the manifestation of early skills. As Spelke and Newport (1998) wrote, asking if knowledge is innate amounts to asking if knowledge is independent of learning and *not* to asking if knowledge is independent of environmental influences: structures of knowledge are intrinsic to the organism and the environment only reveals these structures (Piattelli-Palmarini 1980). This assumption is reflected by two strong nativist positions, respectively initiated by Chomsky (1959), (1965) and Fodor (1983), (1985): the representations and the inferences allowed by representations are innate (Baillargeon 1987; Wynn 1992); the constraints that organize the knowledge of infants are also innate (Keil 1981; Spelke et al. 1992). This second position presupposes, of course, the existence of *representational nativism*.

3. Early Skills without Learning?

To justify the credibility of innate cognitive structures, the advocates of nativist approaches referred to evolutionary approaches and applied to the development of cognition Darwin's arguments concerning the development of species. The logic of the argument is as follows. Psychologists have the difficult task of studying cognition in biological organisms. However, biological organisms are the product of evolutionary forces. This leads to the following conclusion: if the mind is what the brain produces or, rather, if our cognitive and emotional functions are instances of neurobiological factors, then these functions are unavoidably the product of forces related to the evolution of species (Cosmides and Tooby (1994)). For many years, evolutionists considered the relationship between phylogeny and ontogeny in the mode of repetition (Gould 1977). A modern revision of the relationship between phylogeny and ontogeny is proposed by evolutionary psychology. Evolutionary psychology is the application of knowledge and theories of psychology to the understanding of human evolution. From an evolutionary perspective, the behavior is analyzed in terms of cost and benefit in reference to potential for adaptation. This approach was first developed by sociobiology in order to identify human behaviors that have become steady strategies during evolution in a particular environment. In recent decades, a new form of evolutionary psychology has emerged: this approach extends the previous principle to any form of behavior. The mental structures of the modern mind could then be explained in terms of evolutionary adaptation (Tooby and Cosmides 1992; Pinker 1997).