

# Chapter 16

## The Anticipatory Brain: Two Approaches

Mark H. Bickhard

**Abstract** It is becoming increasingly accepted that some form of anticipation is central to the functioning of the brain. But modeling such anticipation has been in several forms concerning what is anticipated, whether and how such ‘anticipation’ can be normative in the sense of possibly being wrong, the nature of the anticipatory processes and how they are realized in the brain, etc. Here I outline two such approaches – the Predictive Brain approach and the Interactivist approach – and undertake a critical comparison and contrast.

**Keywords** Predictive brain • Interactivism • Free energy • Anticipatory brain • Brain models

There are at least two approaches to modeling brain functioning that make central use of notions such as anticipation, expectation, and prediction. I will argue, however, that they diverge at very fundamental levels. The first approach that I will address is a broad programme with multiple variants and various names, such as the Predictive Brain, the Bayesian Brain, and the Free-Energy Principle (e.g., Clark 2013; Friston and Stephan 2007). The second is the interactivist model (e.g., Bickhard 2009a, b).

### 16.1 The Predictive Brain

Predictive brain models have developed within the tradition of Helmholtz’s notion of inference from perceptual sensations to representations of the environment (Clark 2013). This sensory empiricist tradition, in turn, extends to the classical Greeks.

A relatively recent major step forward in this tradition was the development of analysis by synthesis models, in which sensory inputs are analyzed via some sort

---

M.H. Bickhard (✉)  
Lehigh University, Bethlehem, PA, USA  
e-mail: [mhb0@Lehigh.EDU](mailto:mhb0@Lehigh.EDU)

of synthesis of those inputs (MacKay 1956, 1969; Neisser 1967). This move in effect substitutes abductive processes for the deductive and inductive processes that Helmholtz depended on.

One important consequence of this development was the introduction of an active brain, one that is ongoingly engaged in the constructive processes of abductive prediction and explanation. This is in contrast to classical reflex arc and other forms of passive, reactive models.

The predictions of an analysis by synthesis process must be in some way compared to the actual pattern and flow of inputs, and this requires that some signals be generated that can be compared in this manner – either successfully or unsuccessfully matching the actual inputs. A natural next elaboration of such models is to recognize that some such internally generated signals might evoke muscle activity, thus action, and thus *induce* the sensory inputs that are being “predicted”, not just receive them. This yields models such as *Behavior: The Control of Perception* (Powers 1973). In such models *interaction* becomes central, not just reception.

Computational and information processing models were dominant when analysis by synthesis was introduced, but the general form of predictive synthesis could take other forms as well. One such alternative form that has become widely accepted is that of Bayesian inference. Bayesian inference starts with a prior probability distribution over some space of hypotheses and, taking new data into account, modifies that distribution into a posterior distribution. Such a posterior probability distribution, in turn, can constitute the predictions concerning the (probability distributions) of input patterns and flows. A probability distribution with density concentrated at a point can be taken to predict a single particular input. In cases of real number parameterized spaces of hypotheses, the distributions are characterized in terms of their statistical parameters.<sup>1</sup>

The introduction of Bayesian forms of prediction brings with it the reliance on available prior probability distributions, with important additional model requirements and resources. In standard Bayesian statistics, such priors are often treated as strengths of prior beliefs about the available hypotheses, which are then modified with respect to data. But such prior distributions can also be made dependent on previous data and the distribution manifested in such prior “experience”, thus yielding “empirical” prior probability distributions as basis for Bayesian modification into posterior distributions.

One possible source for such prior distributions would be a higher level of Bayesian inference whose posterior distribution parameters could be input as priors to a lower Bayesian process. An iteration of this modeling step yields a hierarchy of Bayesian process layers, with each layer receiving errors of prediction – that which is to be predicted by this layer – *from* below and sending predictions *to* the layer below, with the bottom layer predicting sensory inputs and generating prediction

---

<sup>1</sup>Sufficient statistics (Friston et al. 2009).

errors relative to those errors to the layer above.<sup>2</sup> Such a hierarchy models multiple layers each of which attempts to predict the inputs from below, generated as errors of prediction at the level of that layer below, thus attempting to account for what lower layers have not been successful in accounting for.

Friston has proposed an integrating framework for such hierarchical predictive models (Friston 2013; Friston and Stephan 2007). In this framework, a statistical notion of “free-energy” substitutes for Bayesian parameter estimates. This is in part a practically motivated substitution, but it also has theoretical implications. Practically, minimal error Bayes estimates are difficult to compute; free-energy is proven to constitute an upper bound for Bayes error, so minimizing free-energy entails (approximately) minimizing Bayes error, and free-energy minimization is a tractable computation (Friston et al. 2012a; Kiebel et al. 2008). Such minimizations are computed, in general, via some sort of descent iterations settling at an available minimum. Minimizing free-energy, thus, accomplishes minimizing prediction error, and thus minimizes “surprise”.

Within this framework, Friston proposes that some “predictions” are of proprioceptive inputs, which, in turn, induce muscle activity that yields those inputs, thus inducing action, which in turn induces sensory inputs. This extends the basic hierarchical model to include interactive *inducing* of inputs along with the basic *reception* of inputs.

Within this sensory-action framework, Friston proposes a theoretical modeling use of the reliance on Bayesian prior probability distributions. The hierarchy of Bayesian layers cannot extend unboundedly: there must be some finite number of layers and the highest layer must take its prior distributions from some source other than a still higher layer. It is proposed that they are innate. These highest level prior probability distributions, also called beliefs or expectations, can then be understood as inducing action that yields various kinds of sensory patterns and flows. In this perspective, the induction of action via “expectations” becomes as or more important than simple prediction of input.

In fact, Friston claims that all considerations of properties such as utility, cost, value, and so on can be folded into what counts as surprise relative to these highest level expectations (or beliefs or prior probability distributions – sometimes called hyperpriors). This is because an organism “predicts” what is “expected”, and induces those “expected” inputs via action. What is “expected” in innate hyperpriors (so it is claimed) is flows of inputs that are *viable* for the organism, and, thus, that capture or fold-in considerations of utility and value. The argument for this is that, if ancestral organisms did not expect, thus induce, viable flows of action and inputs, they would have died out, thus extant organisms *must* “expect” such viable flows (Kiebel et al. 2008; Friston et al. 2012b).

---

<sup>2</sup>For an early version of such hierarchical prediction, see Tani and Nolfi (1999).

### 16.1.1 *Some Problems*

This is an ambitious, sophisticated, and impressive modeling framework. Nevertheless, I contend that it suffers from significant problems at every level of its structure.

At the broadest level, this programme is a variant of sensory empiricism, based on sensory encodings as inputs. This is a dominant approach to perception, and has been for millennia. Helmholtz is a major figure in this tradition. There are three ‘obvious’ issues involved in this construal of perception: (1) What is the nature and origin of the inputs upon which the inferences are based? (2) What is the nature of the inference process(es)? And (3) What is the nature and origin of the representations by which the inferred world is characterized?

Classical answers would have it that the inputs are sensations, the inference processes are deduction and induction, and the representations of the wider world are aggregations or organizations of the sensations. In other words, the process is one of inferring world encodings – objects, properties, etc. – on the basis of input encodings. This is basic empiricist epistemology.<sup>3</sup>

Despite millennia of trying, no one has been able to model how causal inputs can generate representational sensory encodings. This is the problem of sensations, sense data, and other basically equivalent terms: the “transduction” of representation from causal inputs (Fodor and Pylyshyn 1981; Bickhard and Richie 1983).<sup>4</sup>

The tradition faces serious problems even if the problem of the mysterious transduction of sensory encodings is overlooked. In particular, not only has no successful model of transduction ever been offered, there has also never been a successful model of how higher level representations could be generated on the basis of such sensory encodings. How, for example, can representations of chair, bachelor, the number three, triangle, or justice be generated or constructed out of sensory encodings?<sup>5</sup>

Bayesian hierarchical models would, in fact, seem to have weaker resources in this regard than classical sensory empiricism. Classical empiricism could attempt to model such representations with various collections and structures of sensory encodings, while the Bayesian hierarchy is limited to parameters of probability distributions over spaces of parameters of probability distributions over . . . over patterns and flows of inputs.

---

<sup>3</sup>E.g., Locke, Hume, Russell (in some incarnations), Fodor, much of contemporary literature, and even Aristotle’s signet ring impressing its form into wax.

<sup>4</sup>“Transduction, remember, is the function that Descartes assigned to the pineal gland” (Haugeland 1998, p. 223).

<sup>5</sup>Or perhaps they’re independently innate? This issue is alive and well in contemporary work: in child development, for example, a fundamental question is whether or not it is possible to construct, say, object encodings or number representations, out of sensory encodings. Some say yes, and some say that such higher-level encoding representations must be innate. Ultimately, neither stance is successful (Allen and Bickhard 2013a, b).

There is, thus, no account of the nature or origins of sensory representations in these models nor of representations of the world in general.<sup>6</sup> Still further, there are in-principle arguments that such accounts are not possible (e.g., Bickhard 2009a). Analysis by synthesis introduces abduction along with deduction, but does not offer anything new regarding an account of representation. Taking into account that action can influence or induce or ‘control’ perceptual inputs is a major step forward, but, yet again, this offers nothing new regarding the nature of perceptual (or any other) representations.

One aspect of this problem is to note that, just as there is no account of representation, so also is there no account of representational error. There is some sort of comparison of ‘predictions’ and inputs, and an iterative descent process<sup>7</sup> which tends to minimize overall discrepancies in such comparisons, but such discrepancies constitute error only for an external observer or designer who interprets those inputs as representations and who interprets those discrepancies as error. The machine or organism that iterates in some descent process has no emergent normativity at all, including that of representation and representational error. The sense in which the model might be considered to be a model in which (sensory) representational error is minimized (or any other kind of error) is limited to that of the interpretive perspective of a designer or some other kind of external interpreter.<sup>8</sup>

Yet another aspect of this designer or interpreter dependence of Bayesian hierarchical models is that there is no account of the spaces over which the probability distributions are distributed. These spaces are required and assumed in the nature of Bayesian estimation; their nature and origin is not addressed.

### 16.1.2 *Free-Energy Problems*

A central claim for the free-energy model is that it incorporates normative considerations, such as cost or utility, into the higher level expectations – hyperpriors – without having to calculate with respect to such considerations: the organism

---

<sup>6</sup>This literature proceeds within a background assumption of semantic information models, conflating technical covariation information with representational (about) information, without ever addressing this assumption. It is, nevertheless, evident everywhere, including Clark (2013).

<sup>7</sup>With respect to some underlying metric on the underlying space.

<sup>8</sup>Block and Siegel (2013) suggest that a better term than “error” might be “discrepancy”, but this too suggests a normative standard from which the “predictions” are “discrepant”. “Difference” is more neutral in this regard: overall, the dynamics of such a system settles into a minimization of such differences. There are no “errors” (Bickhard and Terveen 1995). There are multiple similar abuses of language in this literature, such as “error”, “representation”, “cause”, “belief”, “expectation”, “describe”, etc. none of which (in this literature) refer to anything like the phenomena that such words are generally taken to refer to. Nevertheless, they leave the suggestion, without argument, that they do constitute models for the phenomena at issue (McDermott 1981).

‘simply’ expects to stay in viable zones of its states, and induces actions that keep it in those zones.

Friston and Stephan (2007) attempt to account for normativity in the model via:

The basic premise we start with is that biological systems must keep  $\gamma$  [“ $\gamma$  can be regarded as sensory input”] within bounds (i.e. phase-boundaries) through adaptive changes in  $\alpha$ . [“effect[s] of the system on the environment”] Put simply, adaptive systems or agents should minimise unlikely or surprising exchanges with the environment. (Friston and Stephan 2007, p. 425)

The idea is that evolutionary selection will eliminate organisms that fail to keep sensory input within bounds, and, thus, to generate organisms that minimize surprise – minimize expectation failure – with respect to those inputs. There is in fact more than one premise in these two sentences, and they are all questionable, but my focus here is on the normativity issue. The purported connection of this model to normativity is the notion of “keeping inputs within bounds”. This is not defined, though a supposed characterization is given in “i.e., phase-boundaries”. Unfortunately, phase boundaries are part of the infrastructure of all living systems, and changes in phase-boundaries are intrinsic to growth and development. The failure of phase-boundaries to capture what “within bounds” could mean leaves the notion of “not within bounds” hanging as a surrogate for whatever seems not good for the organism from the observer’s (or Friston’s) perspective. This observer subjectivity of normativity for the organism is not a property of normativity that is fundamental in the constitution of the organism itself (Kiebel et al. 2008).

It is not something intrinsic to the nature of the organism per se, but, rather, is a supposed consequence of the effect of prior evolutionary selection. The organism, in this model, is still ‘just’ factually and causally settling in to conditions that minimize discrepancies between internally generated signals and input generated signals. There is still no normativity, thus, no error.

### 16.1.2.1 What Is It to Be a Living Being?

There are allusions to such a model capturing various aspects of self-organization, embodiment, and other properties of living beings, but it offers a strikingly weak characterization of what it is to be alive: tending to stay in a particular set of states.

Friston’s model has at best a normativity of avoiding dissolution (Friston 2012) – keeping things “within bounds” (which is not captured by notions of phase-boundaries). Living systems clearly do have to avoid dissolution, but, with this “selection designed” dissolution-avoidance, the model captures less of the nature of living beings than, for example, enactivist notions of living beings being autopoietic – continuously constructing and reconstructing their components (Maturana and Varela 1980; Varela 1979, 1997; Di Paolo 2005; Weber and Varela 2002). This is a real time dynamic characterization, not simply a result of ancestral evolution.

The free-energy grounding claim is in even stronger contrast with the recognition that living beings have to engage with the *environmental thermodynamic and material relationships* for their own continued existence – they have to maintain those relationships in order to self-maintain their own continued existence, in order to recreate components, for example (Bickhard 2009a). Living beings are *constituted* as processes that are *open systems of ontological necessity*, not that they *just happen* to be open to inputs from the environment (see Friston’s example of a snowflake that is “open” to its environment, Friston and Stephan 2007, p. 423). Living systems are *constituted* as processes that maintain their thermodynamic conditions in sensitivity to changes in environmental conditions (Bickhard 2009a). They *could not* be closed systems, whereas a snowflake, in principle, could. The normativity of living systems, including of their representing, emerges in this necessity to maintain, create, obtain, and exploit the conditions and the material for their own existence (Bickhard 2009a), not just to avoid dissolution.<sup>9</sup>

This property of the self-maintenance of the conditions for ones own continued existence is a much deeper condition than simply “avoiding dissolution because ancestors were selected for doing so”. Selection histories do not suffice: for example, some crystals grow in a certain way because earlier deposition of molecules from solution happened to realize a particular form of crystal structure, and the seed crystal thus created selects further depositions to fit that form – this is a case of historic selection (Bickhard and Campbell 2003), but there is no life and no normativity: neither a molecule nor the crystal are alive, nor is either *in error* if the molecule should happen to deposit in a position or orientation consistent with a different crystalline form. (Evolutionary) selection histories can *create* systems that are alive and with emergent normativity, but neither being alive nor that normativity is *constituted* in having such a selection history.

### 16.1.2.2 Error Minimization

Overlooking problems concerning the nature of normativity, there are seemingly obvious problems for an error minimization stance, such as the “dark room problem” (Clark 2013): If the organism seeks to minimize prediction error with regard to inputs, why doesn’t it just head to a dark room and stay there? Input prediction error would seem to be fully minimized when there are no inputs. There is a seemingly ready answer within the hyperprior framework: the organism *expects* to be in lighted areas, and so turns on the light in order to fulfill that expectation (Friston et al. 2012a; Friston 2013). This is a straightforward extension of the basic notion that all normative considerations are built in to the hyperpriors (Friston et al. 2012b).

---

<sup>9</sup>This connection between cognition and life was at the center of the interactivist model from its inception, e.g., “knowing as explicated above is an intrinsic characteristic of any living system” (Bickhard 1973, p. 8; also Bickhard 1980a, p. 68). This is also a strong intuition in the enactivist framework, but, so I argue, is not so well captured by the definition of autopoiesis.

But there will be exceptions, such as if the organism wants darkness, perhaps in order to sleep or hunt or hide. These exceptions too can be built-in to the hyperpriors, but a problem begins to be discerned: how many exceptions and exceptions to the exceptions have to be built-in in order to accommodate all normative considerations into hyperpriors? How could they all possibly be ‘built-in’? Isn’t such an epicyclic elaboration of hyperpriors a *reductio* of the approach?

Another example would be an animal on a shock grid. When the tone that signals impending shock is heard, why doesn’t the animal ‘simply’ predict pain and stay on the grid? Why would it learn to jump off of the grid? Perhaps absence of pain is built into the hyperpriors? But what about the exceptions of seeking pain inputs, such as with hot peppers? Are those also ‘built-in’ to the hyperpriors?

Hyperprior ‘expectations’ seem to handle normative phenomena only insofar as all relevant normative considerations are already built-in to those hyperpriors. That does not constitute an adequate explanation or model (Gershman and Daw 2012; Roesch et al. 2012). Thus, there is no model of how or why an organism avoids harm or seeks value. It is simply assumed that it does so as a result of training or evolution.<sup>10</sup> For organisms, then, the assumption is that an adequate set of expectations to be able to account for all normative behavior and behavior learning is somehow evolved into the hyperpriors.

### 16.1.2.3 “Built-in” to What?

Friston at times writes as if these highest level expectations are somehow ‘built-in’ to the organization of the whole organism, not just the brain (e.g., Friston et al. 2009). Such a perspective is clearly more powerful than the assumption that they are innate in the brain, but the most obvious manner in which this could be the case is for the organism to be ‘organized’ so that it tends to avoid hunger and pain etc., and tends to seek pleasure and excitement, etc. But such explicit cost and utility considerations and their underlying processes and neural realizations are precisely what is claimed to be obviated by the hyperprior model.

These claims and the mathematics that support them are mostly elaborations of mathematical equivalences to the basic assumption that the organism tends to stay in some set of ‘expected’ states, and that those states will tend to be viable states whose expectations are induced by evolution (Friston 2012). That basic set of ‘expected’ states, then, has to accommodate all normative considerations.

---

<sup>10</sup>But training has to be with respect to some normative criteria, and there are none other than what is built-in to the hyperpriors.

#### 16.1.2.4 Learning

Hierarchical Bayesian prediction models have powerful resources for learning: hierarchies of Bayesian layers, each generating predictions of, and thus accommodating, the errors of prediction from layers below. Each layer can learn *sequences* of the sufficient statistic parameters for probability distributions over the spaces of the layers below, which, in turn, can learn sequences of the layers below them. So the Bayesian layers can collectively learn sequences of sequences of . . . of patterns and flows of inputs (Kiebel et al. 2008) (perhaps as induced by action ‘predictions’).

The highest level of such expectations, hyperpriors, is at times characterized as a set-point (e.g., Friston et al. 2012b), and the overall hierarchy does have a flavor of a servomechanism hierarchy, in which higher levels send (sequences) of goals to lower levels. But learning sequences of sequences cannot capture the unbounded spaces of possible interactions that even a simple feedback servomechanism can manifest. The Bayesian hierarchy can generalize beyond actually experienced flows of interaction, but only in terms of sequences of sequences related by the descent processes used to calculate minimum free-energy solutions.

It is crucial to note, in this regard, that the underlying spaces upon which these calculations take place, as well as the metric organizations on those spaces that are necessary for any generalizations to take place – for descent calculations to proceed – must be already available for the calculations to take place. That is, they must be innate. The Bayesian processes offer no way in which new such spaces or new such metrics can be generated, so the spaces and their metrics must not only be innate, they are fixed (e.g., Kiebel et al. 2008): new layers cannot be generated and new metrics, or any other kind of topological or metric organization, cannot be organized. New layers of generalization, thus, as well as the metrics and topologies of generalization – and thus new kinds of generalization – cannot be learned. This holds at all layers of a Bayesian hierarchy, not just the highest hyperprior layer. Yet we know that analogy and metaphor, for example, can induce new cognitive organizations, can re-organize similarity spaces – that is, can change *via learning* the topological and metric organizations involved (Gentner and Rattermann 1991; Gentner and Jeziorski 1993; Medin et al. 1993).

It would be possible for one layer to switch a lower layer from one space to another, perhaps with a different metric (Friston et al. 2009), but, again, only if the spaces to be switched among are already innately available in the lower layer. Again, this is an example in which everything has to be already pre-prepared for whatever will be needed.

#### 16.1.2.5 CNS Architecture

Friston claims that the Bayesian layers of a hierarchical Bayesian prediction process are realized in brain architecture, and, thus, that hierarchical Bayesian prediction layers constitute the basic functional architecture of the brain (Friston and Stephan 2007; Kiebel et al. 2008; Adams et al. 2012). Such a characterization

of CNS architecture makes partial sense for some cortical domains in early sensory processing, especially for the visual system, though it is not (yet) clear that the microfunctioning of those ‘layers’ fits the model (Clark 2013; Friston 2008), and it is argued that motor functioning can also be (partially) accommodated (Adams et al. 2012).

With respect to larger and more general considerations, however, the hierarchical Bayesian model does not fit well. For example, there are large parts and regions of the brain that do not have a hierarchical organization, such as multiple node loops involving cortex and sub-cortical regions (Doya 1999; Koziol and Budding 2009; Nieuwenhuys 2001)<sup>11</sup> – such as prefrontal to striatum to thalamus to prefrontal, or prefrontal to cerebellum to thalamus to prefrontal, and so on. Multiple-node loops abound and are not consonant with a simple hierarchy.

Furthermore, there are multifarious characteristics of brain functioning that are not touched upon by the neural network hierarchical Bayesian model. These include silent neurons that rarely or never fire, gap junctions, intrinsic oscillatory activity (not just iterative re-entrance) at both neural and circuit levels, wide-spread release of neuro-modulators that can functionally reorganize activity (Doya 2002; Marder and Thirumalai 2002; Marder 2012), emotional processes and the limbic portions of the brain that underlie them, episodic and biographical memory and learning, and so on. Most of what we know about the brain is either not accounted for by these models or is flatly inconsistent with them.

#### 16.1.2.6 So, Hierarchical Bayes?

Hierarchical Bayes, thus, is simply not powerful enough to account for multiple phenomena of normative activity, of learning and development, of neural functioning at both micro- and macro-levels, and how any of these are related to the nature of being alive.

There are fundamental problems resident in the basic framework of sensory encoding assumptions about the nature of representation; problems concerning the nature of representation are not addressed in these models. Additions to the sensory inference model in the form of abductive anticipation, action as inducing sensory inputs, hierarchical Bayesian processes of prediction, and free-energy claims to account for normative phenomena and for brain functional architecture each add additional power to the resources of the overall model, but do not resolve any of the difficulties inherited from earlier model innovations, and each introduces new problems of its own.

I will argue, nevertheless, that there are crucial insights in these models that need to be maintained – in particular, the importance of anticipation and of *interaction*, not just inputs and action – and I outline a model that arguably does so without encountering the problems of free-energy hierarchical Bayes.

---

<sup>11</sup>In spite of brief mention of such architectures in Adams et al. (2012).

## 16.2 The Interactivist Model

Predictive encoding models focus on prediction of inputs at the sensory interface. In contrast, consider the possibility that it is *flow of interaction* that is anticipated, not the inputs (or outputs) that participate in that interaction. A perspective on this possibility can be found at the root of the interactivist model:

Consider two Moore machines [abstract finite state machines with outputs] arranged so that the outputs of each one serve as the inputs of the other. Consider one of the Moore machines as a system and the other as its environment, and let the system have the initial and final state selections that make it a recognizer.

The system can thus recognize input strings in the standard sense in automata theory [a recognizer “recognizes” strings of inputs that move it from its initial state to one of its final states]. In this interactive configuration, however, an input string corresponds to – is generated by – a state transition sequence in the environment. The set of recognizable input strings thus corresponds to the particular set of state sequences in the environment that could generate them. The recognition, or knowing, relationship is thus extended from inputs to situations and conditions in the environment.

Furthermore, during an interaction, the environment is receiving outputs from the system – and it is these outputs from the system that induce the environmental state transitions that generate the inputs to the system that the system either recognizes or doesn’t. Thus the ‘recognition’ process is no longer strictly passive – the ‘recognized’ strings are induced from the environment by the system’s own outputs. In fact, the interaction doesn’t need to be viewed as a recognition process at all. It is equally as much a construction or transformation process – constructing the situations and conditions corresponding to the last state of a ‘recognizable’ environmental state sequence – or at least a detection process – detecting an initial state of a ‘recognizable’ environmental state sequence – and so on.

The system need not be thought of as a single undifferentiated recognizer. It could be, for example, a collection of recognizers connected to each other, say, with the final states of one attached to the initial state of another. Such connections could induce functional relationships among the recognizers, such as one testing for the appropriate conditions for another to begin, or a servomechanism being used to create a subcondition for another process to proceed, etc. (Bickhard 1973, pp. 21–22; also in Bickhard 1980a, pp. 75–76)

Here we have a formalization within abstract machine theory of the idea of recognizing input strings, and of inducing those “recognizable” input strings from the environment. This sounds a bit like the prediction and control of inputs of action oriented predictive processing (Clark 2013). And it is.

But the contrasts are crucial. First, the inputs in the interactivist model are not mis-construed as representational; they are not supposed to be transductions-into-representations. Instead, they are simply registrations that move the machine in its state transition diagram. Representation emerges in the implicit detection or presupposition relationships between the induction of recognizable strings and the environmental conditions that support such recognizable string induction. This is a fundamentally different conception of representing, and one that does not encounter the myriad problems of encodingisms. *It is the anticipating that is*

*representational* – truth valued – *not the inputs*. This is a pragmatic, action based model of representation, and it is much stronger than any encodingist model.<sup>12</sup>

Note in particular that if there is a connection from one recognizer to another, then the ‘anticipation’ that some environment recognized or transformed by the first will also be one that is ‘recognizable’ by the second could be false, and could be discovered to be false in virtue of the second recognizer not entering one of its final states (Bickhard 1980a), perhaps, for example, encountering a halting condition instead. Another consequence is that, because inputs are not representational, the representation relationship is not restricted to things that can be ‘represented’ by aggregations and organizations of inputs.

The second difference is that the abstract machine configuration is not in itself normative. There is no normative difference from the machine perspective between ‘successfully’ recognizing or inducing relevant input strings or ‘failing’ to do so. All such processes and outcomes are ‘just’ factual, ‘causal’, processes in the machine(s) and environment. Accounting for normativity requires further fundamentally thermodynamic considerations (Bickhard 1993, 2009a) that are not present in the abstract machine model. Predictive encoding models, in contrast, treat the inputs as transductions and expectations as normative, when in fact they too involve nothing more than factual, causal processes. With no normativity, there is no error; predictive encoding models settle into conditions in which they minimize discrepancies between inputs and internally generated signals that are matched against those inputs. They are a complicated form of unsupervised “learning” (Bickhard and Terveen 1995), but do not transcend the basic limitations of such models.

### 16.2.1 *What and Where of Anticipation*

Predictive brain models focus on anticipating or inducing global inputs, while the interactivist model focuses on anticipating interactive flow. Anticipating interactive flow has local aspects, as well as global aspects. In particular, local domains of the brain have to successfully anticipate their own local flow of process in order for the whole brain to successfully anticipate its global, thus interactive, processes. In shifting from anticipation of inputs to anticipation of interactive processing, thus, we also shift from global perspectives to local perspectives.

In particular, each local domain of the brain engages in anticipatory processes concerning its own local near future processes. This is a primary focus of the interactivist model. Local anticipatory processes are realized in local “set-ups” of preparation for particular ranges of potential further process flow – a “microgenesis”

---

<sup>12</sup>For discussions of action based anticipatory models, see, for example, Bickhard (1980a, b, 1993, 2009a, b), Bickhard and Richie (1983), Bickhard and Terveen (1995), Buisson (2004), Pezzulo (2008), Pezzulo et al. (2013).

of preparation of local conditions for further processes (Bickhard 2009c, 2015a, b, [in preparation](#)). Such microgenesis, thus, anticipates that the future will remain consistent with what has been microgenetically set-up for.

Of central importance is that microgenetic anticipation can be correct or not correct: true or false. This is the locus of the emergence of representational truth value. Much more needs to be done to model more complex forms of representation, but truth value is the central barrier to naturalistic models of representation (for further development of the representational model, see, e.g., Bickhard 1980b, 1993, 2009a, 2015a, b, [in preparation](#)).

## 16.2.2 Learning

Learning is a major integrating perspective in this model for brain functioning and brain evolution (Bickhard 2015a, b). Here I will introduce one functional assumption and argue that it suffices to account for multiple forms of learning.

That assumption is that successful microgenetic anticipation yields consolidation of those microgenetic processes in those functional circumstances, while unsuccessful microgenetic anticipation yields destabilization. This models a variation and selection process – an evolutionary epistemology – with successful (microgenetic) anticipation as the selection criterion (Campbell 1974; Bickhard and Campbell 2003).

### 16.2.2.1 Habituation

If some part of the nervous system has a simple matching or subtractive relationship with its inputs, then microgenetic anticipation of the processes of input registration will amount to anticipation of those inputs *per se*. This is the basic model for Sokolov habituation (Sokolov 1960). For a simple tone, this requires only the first cochlear nucleus to be able to successfully anticipate.

### 16.2.2.2 Classical Conditioning

If pain inputs are inputs that cannot be habituated, and, furthermore, offer no possibility of successful interaction (to a first approximation), then the only way to successfully interact with a sequence of tone followed by (say) shock is to avoid the shock – jump off of the shock grid. Note that there are no “hyperprior” “expectations” involved.

### 16.2.2.3 Instrumental Conditioning

If the hypothalamus is generating signals as a result of low blood sugar, then the only way to (ultimately) interact with this input is to do something that raises blood sugar.<sup>13</sup> These activities can be enormously complex, as well as contextually and culturally variable, though fussing and crying generally suffice for infants (together with supportive forms of interaction such as rooting and sucking, as well as a supportive environment). In any case, learning will stabilize on forms of interaction that successfully halt or diminish the hypothalamic signals.

### 16.2.2.4 Other Forms of Learning

These forms of learning all involve successful termination or diversion of an input stream. But this is not the only manner in which microgenetic anticipation can be successful. If microgenesis can proceed in a temporal trajectory that ongoingly successfully anticipates the ongoing trajectory of process flow, then that too manifests successful microgenetic anticipation. These forms of microgenetic anticipation underlie forms of learning such as incidental learning, memory, and so on (Bickhard 2006, 2015a, b). I will not elaborate these points here,<sup>14</sup> but will proceed to some further comparisons with predictive brain models.

## 16.2.3 Partial Convergences

Predictive brain models can account for Sokolov habituation very directly: actual inputs are matched by predicted inputs, and the predictions are successful. If the only relevant activity of some part of the brain is to register inputs, then local microgenetic anticipation is extensionally equivalent to input (registration) prediction, and there is a close convergence between interactivist and predictive brain models. They are both models of anticipatory or predictive processes at the center of brain functioning, and the two different senses of anticipation/prediction have a strong convergence for habituation.

The difference between predicting inputs and anticipating local microgenetic processes is subtle in this case, but yields wide divergences for more complex cases. In effect, predictive brain models assimilate *all* brain processes to complex habituation (successful input subtraction) processes.

---

<sup>13</sup>Hunger and eating is much more complex than this, with multiple feedforward and feedback processes, but this captures the basic organization of the phenomenon (Carlson 2013).

<sup>14</sup>This paper is not the occasion to attempt to present the entire model. I have addressed only enough to be able to make some comparisons with predictive brain models.

Classical conditioning begins to manifest a wider divergence. For the interactivist model, jumping off of the shock grid is a successful way to interact with, thus successfully anticipating the interactive process flow, a tone-to-shock input stream: prevent the shock. The free-energy version of the predictive brain models must hypothesize some sort of innate hyperprior “expectation” that the organism will not experience shock or pain, which then yields some sort of steepest descent convergence on a behavior of jumping off the grid.<sup>15</sup> This, as mentioned earlier, is ad-hoc and has difficulty accounting for exceptions such as for eating hot peppers. The interactivist model has a ready resource for accounting for such exceptions: if the organism has learned to interact with some pain inputs, such as from peppers, with ongoing microgenetic anticipation of the processes evoked by those inputs, then that too is successful anticipatory interaction.<sup>16,17</sup>

To reiterate: the focus of predictive brain models is on the global organism interface with sensory inputs, extended to include proprioceptive “inputs” as a way of accommodating action. In these models, representation is somehow based on the inputs as sensory representations, as well as on the innate spaces, metrics, and probability distributions over those spaces as layers of a Bayesian hierarchy are taken into consideration. In the interactive model, *it is the anticipating that is representational*, not the inputs. The inputs, as well as outputs, play an essential role in influencing the processes that microgenesis is “attempting” to anticipate, but neither inputs nor outputs per se are representational in this model. Among other consequences is that the model completely transcends the problem of accounting for some sort of transductive interface between organism and world.

---

<sup>15</sup>Note that “steepest descent” processes are not nearly as general as an evolutionary epistemology. This is another aspect of the fact that the Bayesian models require pre-given spaces, metrics on those spaces, and innate distributions (expectations) at least at the highest “hyperprior” level.

<sup>16</sup>And such forms of interaction – e.g., with peppers – will not evoke negative emotional processes, such as fear and anxiety. I will not present the interactivist model of emotions here, but wish to point out that they too are involved in successful microgenetic anticipation (Bickhard 2000).

<sup>17</sup>Insofar as the highest level hyperpriors are “built-in” to the whole organism, not just into the nervous system, it might be claimed that such properties of pain inputs are what constitute the relevant hyperprior(s) for pain. But the interactivist model for pain and for learning with respect to pain is a selection model, a cost or utility or normative model, which – as mentioned earlier – is precisely what Bayesian hyperpriors are supposed to obviate the need for. Thus, to make such a claim contradicts the supposed ability of the Bayesian hierarchical predictive brain model to do without explicit cost or norm considerations.

## 16.2.4 *Functional Processes in the Brain*

At this point, I will present a brief overview of how interactive processes are realized in the brain.<sup>18</sup> As will be seen, this view accommodates multiple phenomena that are known about brain functioning, but are at best anomalous for standard views. The overview will be in two parts: a focus on micro-functioning and a focus on macro-functioning.

### 16.2.4.1 *Micro-functioning*

The interactivist model focuses on interaction as the fundamental locus for the emergence of representation and cognition (Bickhard 2009a).<sup>19</sup> This ranges from interactions of organisms with their environments to interactions of local brain regions with their surrounds, including other parts of the brain, the body, and the environment.

One consequence of a focus on interaction is the necessity of accounting for timing: successful interactions must have the right timing relationships with whatever is being interacted with in order for those interactions to be successful.<sup>20</sup> Timing, in turn, requires clocks, but clocks are “just” oscillators and the easy way to accommodate the need for oscillators is for all functional processes to be realized as oscillatory processes modulating each other. This modeling framework is *at least as powerful* as Turing machines: a limit case of modulation is for one process to switch another on or off, and Turing machines can be constructed out of switches. It is *more powerful* than Turing machine theory in that it has inherent timing.

If we examine brain processes, multifarious forms of endogenously active oscillatory and modulatory relationships is precisely what we find. These range from the spatially small and temporally fast, such as gap junctions, to large and slower, such as volume transmitters and astrocyte influences:

- silent neurons that rarely or never fire, but that do carry slow potential waves (Bullock 1981; Fuxe and Agnati 1991; Haag and Borst 1998; Roberts and Bush 1981);
- volume transmitters, released into intercellular regions and diffused throughout populations of neurons rather than being constrained to a synaptic cleft (Agnati et al. 1992, 2000); such neuromodulators can reconfigure the functional prop-

---

<sup>18</sup>See, for example, Bickhard (1997, 2015a, b, in preparation; Bickhard and Campbell 1996; Bickhard and Terveen 1995).

<sup>19</sup>In thus focusing on action and interaction, the interactivist model is in strong convergence with Piaget and with other pragmatist influenced models (Bickhard 2006, 2009a). (There is in fact an intellectual descent from Peirce and James through Baldwin to Piaget.) There are also interesting convergences of this model with Dewey.

<sup>20</sup>Timing goes beyond sequence, and, thus, goes beyond Turing machine theory and equivalents (Bickhard and Richie 1983; Bickhard and Terveen 1995).

erties of “circuits” and even reconfigure functional connectivity (Marder and Thirumalai 2002; Marder 2012);

- gaseous transmitter substances, such as NO, that diffuse without constraint from synapses and cell walls (e.g., Brann et al. 1997);
- gap junctions, that function extremely fast and without any transmitter substance (Dowling 1992; Hall 1992; Nauta and Feirtag 1986);
- neurons, and neural circuits, that have resonance frequencies, and, thus, can selectively respond to modulatory influences with the “right” carrier frequencies (Izhikevich 2001, 2002, 2007);
- astrocytes that<sup>21</sup>:
  - have neurotransmitter receptors,
  - secrete neurotransmitters,
  - modulate synaptogenesis,
  - modulate synapses with respect to the degree to which they function as volume transmission synapses,
  - create enclosed “bubbles” within which they control the local environment within which neurons interact with each other,
  - carry calcium waves across populations of astrocytes via gap junctions.<sup>22</sup>

Thus we find a vast spatial and temporal range of oscillatory and modulatory processes, *all* of which are anomalous from any kind of passive threshold switch or connectionist node modeling perspective.<sup>23</sup>

Crucial to my current purposes is that the spatially larger and (thus) temporally slower processes influence the smaller faster processes – which include the classical synaptic influences – by modulating the local environments, such as ion and transmitter concentrations.<sup>24,25</sup> The slower time scales of such processes imply that they, in effect, set parameters for the faster processes. At the faster scales, these parameters are approximately constant, though they undergo their own trajectories of change at those slower scales. Parameter setting for endogenously active dynamic

---

<sup>21</sup>The literature on astrocytes has expanded dramatically in recent years: e.g., Bushong et al. 2004; Chvátal and Syková 2000; Hertz and Zielker 2004; Nedergaard et al. 2003; Newman 2003; Perea and Araque 2007; Ransom et al. 2003; Slezak and Pfeifer 2003; Verkhratsky and Butt 2007; Viggiano et al. 2000.

<sup>22</sup>This list from Bickhard (2015a, b). It is *not* an exhaustive list of the multifarious forms of functioning in the brain.

<sup>23</sup>For a model that also addresses some of these scale phenomena, see, e.g., Freeman (2005; Freeman et al. 2012).

<sup>24</sup>For a discussion of relatively local volume transmitter influences, often characterized as neuromodulators, see Marder (2012; Marder and Thirumalai 2002). This is ‘just’ one class of larger scale, slower forms of modulation.

<sup>25</sup>Kiebel et al. (2008) discuss differential times scales involved in the Bayesian hierarchical model, but, in that model, the time scale differences arise because differing sequences being tracked in the environment change on differing time scales, not because of any differences at the neural and glial level (Bickhard 2015a, b).

processes is the dynamic equivalent of programming for discrete, one step at a time, computational models. The larger, slower, processes, thus, “program” the faster processes: they “set them up” for what those faster processes will be doing in the near future. They constitute microgenesis processes, and, thus, microgenetic anticipation processes.

Note that, just examining what we know about brain functioning, we find these larger, slower processes that set parameters for smaller, faster processes. That is, we find microgenetic “programming”. This is intrinsically anticipative, and, thus, intrinsically has truth value – and we find that the interactive model of representation entails brain properties that we in fact find, and, in reverse, the range of brain processes entail the kind of anticipative processes that constitute the interactive model of representation. This reciprocal entailment is a strong consilience.

Local processes, then, especially in the cortex (Bickhard 2015a), realize the sort of microgenetic anticipatory processes that constitute emergent representation. What modulates those local processes?

#### 16.2.4.2 Macro-functioning

Local processes are modulated by more global processes. These too will be oscillatory/modulatory processes engaged in reciprocal projections among cortical regions, reciprocal projections between thalamus and cortex, multi-node loops involving subcortical regions, such as prefrontal to striatum to thalamus to prefrontal, and so on. Modulation relationships among oscillatory processes are not only inherent in single cell and local processes, but also in the general two and more node loops that make up macro-brain architecture (Koziol and Budding 2009). Differing loops in this architecture will manifest differing sorts of influence on the rest of the brain, such as organization of interaction, including sensory interaction, apperception of the situation in which the organism is located, conditions of dynamic uncertainty concerning what to do next, and so on.<sup>26</sup>

One crucial issue is how the brain arrives at any sort of functional coherence in its activities. How does it achieve functional coherence in the manner in which it engages in its multiple macro-level interactions, and in which it modulates myriad local microgenesis processes? One seemingly obvious answer might be that some central executive controls what the rest of the brain is to do, likely the prefrontal

---

<sup>26</sup>For more specificity concerning such macro-functional considerations, see Bickhard (2015a, b, in preparation). For the general model of perception, apperception, and so on, see Bickhard and Richie (1983; Bickhard 2009a). The model of perceiving offered has strong convergences with Gibson (1966, 1977, 1979), but also some important divergences (Bickhard and Richie 1983). A partial convergence with the model of interactive knowledge of the situation is found in Gross et al. (1999).

cortex. But this “answer” begins a regress of executive decision making: how does the prefrontal cortex achieve its own functional coherence – how does it decide what to do?<sup>27</sup>

Each process in the brain will tend to recruit other processes into modes that yield overall successful interaction and anticipation of successful interaction. In that sense, each process is competing with other processes to generate a self-organization of brain activity into a functioning form that responds to overall interactive success. Brain regions that are not recruited in such self-organizing activity will simultaneously be induced to engage in learning that stabilizes when the region *can* participate in such self-organization, and it will tend to induce such learning changes in other brain processes so that they will tend to generate successful processes in accordance with that local region – that is, the general nature of the processes of competitive recruitment that yields self-organization.

The prefrontal cortex is ideally suited to facilitate such self-organizing integrative processes, but does not have to function as a supreme executive to do so. Instead, it is a locus at which many functional loops converge in such a way that they can participate in global self-organization.<sup>28</sup>

So, multiple domains of the brain are active in differentiating differing aspects of the internal and external situation. Each competes to recruit other domains to interact with its dynamics, thus, with whatever it is differentiating. Striatum loops, cerebellum loops, limbic loops, etc. all specialized for recruiting for special aspects of the situation (Koziol and Budding 2009) – thus (when successful)<sup>29</sup> inducing a kind of self-organization of macro-functioning. This is in strong contrast to the Bayesian brain model of a fixed hierarchy of layers.<sup>30</sup>

## 16.3 Conclusion

Interactive anticipation is the central nature of the emergence of representation. But it is the anticipating that is representational – truth valued – not the inputs (or outputs) that influence internal processes. Global sensory empiricism is the wrong framework for modeling brain anticipatory processes. The brain overall, and each

---

<sup>27</sup>Certainly not via some set of fixed innate hyperpriors.

<sup>28</sup>There is, of course, no guarantee that such self-organization will (fully) succeed at any particular time or in any particular situation.

<sup>29</sup>Lack of coherence is certainly possible, and it can also be functional (in several ways) for the brain to engage in chaotic processes. For example, chaotic processes can be a baseline form of process from which functional attractor landscapes can be induced and controlled (Freeman 1995, 2000a, b; Freeman and Barrie 1994; Bickhard 2008).

<sup>30</sup>Note also that in the Bayesian brain model, the reciprocal projections among various cortical regions are supposed to be engaging in descent iterations, not oscillations (Friston et al. 2012b).

local domain of the brain, engage in “attempting” to anticipate future process, via interaction and microgenesis, and microgenetic recruitment of other domains.

The interactive model and various predictive brain models converge on the centrality of anticipation, but strongly diverge in what that means and how it is manifest in brain functional activity.

## References

- Adams, R. A., Shipp, S., & Friston, K. J. (2012). Predictions not commands: Active inference in the motor system. *Brain Structure and Function*. doi:[10.1007/s00429-012-0475-5](https://doi.org/10.1007/s00429-012-0475-5).
- Agnati, L. F., Bjelke, B., & Fuxe, K. (1992). Volume transmission in the brain. *American Scientist*, 80(4), 362–373.
- Agnati, L. F., Fuxe, K., Nicholson, C., & Syková, E. (2000). *Volume transmission revisited* (Progress in brain research, Vol. 125). Amsterdam: Elsevier.
- Allen, J. W. P., & Bickhard, M. H. (2013a). Stepping off the pendulum: Why only an action-based approach can transcend the nativist-empiricist debate. *Cognitive Development*, 28, 96–133.
- Allen, J. W. P., & Bickhard, M. H. (2013b). The pendulum still swings. *Cognitive Development*, 28, 164–174.
- Bickhard, M. H. (1973). *A model of developmental and psychological processes*. Ph.D. Dissertation, University of Chicago.
- Bickhard, M. H. (1980a). A model of developmental and psychological processes. *Genetic Psychology Monographs*, 102, 61–116.
- Bickhard, M. H. (1980b). *Cognition, convention, and communication*. New York: Praeger Publishers.
- Bickhard, M. H. (1993). Representational content in humans and machines. *Journal of Experimental & Theoretical Artificial Intelligence*, 5, 285–333.
- Bickhard, M. H. (1997). Cognitive representation in the brain. In R. Dulbecco (Ed.), *Encyclopedia of human biology* (2nd ed., pp. 865–876). San Diego: Academic.
- Bickhard, M. H. (2000). Motivation and emotion: An interactive process model. In R. D. Ellis & N. Newton (Eds.), *The caldron of consciousness* (pp. 161–178). Amsterdam/Philadelphia: J. Benjamins.
- Bickhard, M. H. (2006). Developmental normativity and normative development. In L. Smith & J. Voneche (Eds.), *Norms in human development* (pp. 57–76). Cambridge: Cambridge University Press.
- Bickhard, M. H. (2008, May 22–23). The microgenetic dynamics of cortical attractor landscapes. *Workshop on Dynamics in and of Attractor Landscapes*, Parmenides Foundation, Isola d’Elba, Italy.
- Bickhard, M. H. (2009a). The interactivist model. *Synthese*, 166(3), 547–591.
- Bickhard, M. H. (2009b). Interactivism. In J. Symons & P. Calvo (Eds.), *The routledge companion to philosophy of psychology* (pp. 346–359). London: Routledge.
- Bickhard, M. H. (2009c). The biological foundations of cognitive science. *New Ideas in Psychology*, 27, 75–84.
- Bickhard, M. H. (2015a). Toward a model of functional brain processes I: Central nervous system functional micro-architecture. *Axiomathes*. doi:[10.1007/s10516-015-9275-x](https://doi.org/10.1007/s10516-015-9275-x).
- Bickhard, M. H. (2015b). Toward a model of functional brain processes II: Central nervous system functional macro-architecture. *Axiomathes*. doi:[10.1007/s10516-015-9276-9](https://doi.org/10.1007/s10516-015-9276-9).
- Bickhard, M. H. (in preparation). *The whole person: Toward a naturalism of persons – Contributions to an ontological psychology*.
- Bickhard, M. H., & Campbell, R. L. (1996). Topologies of learning and development. *New Ideas in Psychology*, 14(2), 111–156.

- Bickhard, M. H., & Campbell, D. T. (2003). Variations in variation and selection: The ubiquity of the variation-and-selective retention ratchet in emergent organizational complexity. *Foundations of Science*, 8(3), 215–282.
- Bickhard, M. H., & Richie, D. M. (1983). *On the nature of representation: A case study of James Gibson's theory of perception*. New York: Praeger Publishers.
- Bickhard, M. H., & Terveen, L. (1995). *Foundational issues in artificial intelligence and cognitive science: Impasse and solution*. Amsterdam: Elsevier Scientific.
- Block, N., & Siegel, S. (2013). Attention and perceptual adaptation. *Behavioral and Brain Sciences*, 36, 205–206.
- Brann, D. W., Ganapathy, K. B., Lamar, C. A., & Mahesh, V. B. (1997). Gaseous transmitters and neuroendocrine regulation. *Neuroendocrinology*, 65, 385–395.
- Buisson, J.-C. (2004). A rhythm recognition computer program to advocate interactivist perception. *Cognitive Science*, 28(1), 75–87.
- Bullock, T. H. (1981). Spikeless neurones: Where do we go from here? In A. Roberts & B. M. H. Bush (Eds.), *Neurons without impulses* (pp. 269–284). Cambridge: Cambridge University Press.
- Bushong, E. A., Martone, M. E., & Ellisman, M. H. (2004). Maturation of astrocyte morphology and the establishment of astrocyte domains during postnatal hippocampal development. *International Journal of Developmental Neuroscience*, 2(2), 73–86.
- Campbell, D. T. (1974). Evolutionary epistemology. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp. 413–463). LaSalle: Open Court.
- Carlson, N. R. (2013). *Physiology of behavior* (11th ed.). Upper Saddle River: Pearson.
- Chvátal, A., & Syková, E. (2000). Glial influence on neuronal signaling. In L. F. Agnati, K. Fuxe, C. Nicholson, & E. Syková (Eds.), *Volume transmission revisited* (Progress in brain research, Vol. 125, pp. 199–216). Amsterdam: Elsevier.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–253.
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4), 429–452.
- Dowling, J. E. (1992). *Neurons and networks*. Cambridge, MA: Harvard University Press.
- Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia, and the cerebral cortex? *Neural Networks*, 12, 961–974.
- Doya, K. (2002). Metalearning and neuromodulation. *Neural Networks*, 15, 495–506.
- Fodor, J. A., & Pylyshyn, Z. (1981). How direct is visual perception?: Some reflections on Gibson's ecological approach. *Cognition*, 9, 139–196.
- Freeman, W. J. (1995). *Societies of brains*. Mahwah: Erlbaum.
- Freeman, W. J. (2000a). *How brains make up their minds*. New York: Columbia.
- Freeman, W. J. (2000b). *Mesoscopic brain dynamics*. London: Springer.
- Freeman, W. J. (2005). NDN, volume transmission, and self-organization in brain dynamics. *Journal of Integrative Neuroscience*, 4(4), 407–421.
- Freeman, W. J., & Barrie, J. M. (1994). Chaotic oscillations and the genesis of meaning in cerebral cortex. In G. Buzsáki, R. Llinas, W. Singer, A. Berthoz, & Y. Christen (Eds.), *Temporal coding in the brain* (pp. 13–37). Berlin: Springer.
- Freeman, W. J., Livi, R., Obinata, M., & Vitiello, G. (2012). Cortical phase transitions, non-equilibrium thermodynamics and the time-dependent Ginzburg-Landau equation. *International Journal of Modern Physics B*, 26(6), 29 p.
- Friston, K. J. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, 4(11), e1000211. doi:10.1371/journal.pcbi.1000211.
- Friston, K. J. (2012). A free energy principle for biological systems. *Entropy*, 14, 2100–2121. doi:10.3390/e14112100.
- Friston, K. (2013). Active inference and free energy. *Behavioral and Brain Sciences*, 36, J.212–J.213.
- Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, 159, 417–458.

- Friston, K. J., Daunizeau, J., & Kiebel, S. J. (2009). Reinforcement learning or active inference? *PLoS ONE*, *4*(7), e6421. doi:10.1371/journal.pone.0006421.
- Friston, K. J., Adams, R. A., Perrinet, L., & Breakspear, M. (2012a). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology*, *3*, 1–20.
- Friston, K. J., Samothrakis, S., & Montague, R. (2012b). Active inference and agency: Optimal control without cost functions. *Biological Cybernetics*. doi:10.1007/s00422-012-0512-8.
- Fuxe, K., & Agnati, L. F. (1991). Two principal modes of electrochemical communication in the brain: Volume versus wiring transmission. In K. Fuxe & L. F. Agnati (Eds.), *Volume transmission in the brain: Novel mechanisms for neural transmission* (pp. 1–9). New York: Raven.
- Gentner, D., & Jeziorski, M. (1993). The shift from metaphor to analogy in western science. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 447–480). New York: Cambridge University Press.
- Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 225–277). London: Cambridge University Press.
- Gershman, S. J., & Daw, N. D. (2012). Perception, action, and utility: The tangled skein. In M. I. Rabinovich, K. J. Friston, & P. Verona (Eds.), *Principles of brain dynamics: Global state interactions* (pp. 293–312). Cambridge, MA: MIT.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing* (pp. 67–82). Hillsdale: Erlbaum.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gross, H.-M., Heinze, A., Seiler, T., & Stephan, V. (1999). Generative character of perception: A neural architecture for sensorimotor anticipation. *Neural Networks*, *12*, 1101–1129.
- Haag, J., & Borst, A. (1998). Active membrane properties and signal encoding in graded potential neurons. *The Journal of Neuroscience*, *18*(19), 7972–7986.
- Hall, Z. W. (1992). *Molecular neurobiology*. Sunderland: Sinauer.
- Haugeland, J. (1998). *Having thought*. Cambridge, MA: Harvard U. Press.
- Hertz, L., & Zielker, H. R. (2004). Astrocytic control of glutamatergic activity: Astrocytes as stars of the show. *Trends in Neurosciences*, *27*(12), 735–743.
- Izhikevich, E. M. (2001). Resonate and fire neurons. *Neural Networks*, *14*, 883–894.
- Izhikevich, E. M. (2002). Resonance and selective communication via bursts in neurons. *Biosystems*, *67*, 95–102.
- Izhikevich, E. M. (2007). *Dynamical systems in neuroscience*. Cambridge, MA: MIT.
- Kiebel, S. J., Daunizeau, J., & Friston, K. J. (2008). A hierarchy of time-scales and the brain. *PLoS Computational Biology*, *4*(11), e1000209. doi:10.1371/journal.pcbi.1000209.
- Koziol, L. F., & Budding, D. E. (2009). *Subcortical structures and cognition*. New York: Springer.
- MacKay, D. M. (1956). The epistemological problem for automata. In C. E. Shannon & J. McCarthy (Eds.), *Automata studies* (pp. 235–251). Princeton: Princeton University Press.
- MacKay, D. M. (1969). *Information, mechanism and meaning*. Cambridge, MA: MIT Press.
- Marder, E. (2012). Neuromodulation of neuronal circuits: Back to the future. *Neuron*, *76*, 1–11.
- Marder, E., & Thirumalai, V. (2002). Cellular, synaptic and network effects of neuromodulation. *Neural Networks*, *15*, 479–493.
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition*. Dordrecht: Reidel.
- McDermott, D. (1981). Artificial intelligence meets natural stupidity. In J. Haugeland (Ed.), *Mind design* (pp. 143–160). Cambridge, MA: MIT Press.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*, 254–278.
- Nauta, W. J. H., & Feirtag, M. (1986). *Fundamental neuroanatomy*. San Francisco: Freeman.
- Nedergaard, M., Ransom, B., & Goldman, S. A. (2003). New roles for astrocytes: Redefining the functional architecture of the brain. *Trends in Neurosciences*, *26*(10), 523–530.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton.

- Newman, E. A. (2003). New roles for astrocytes: Regulation of synaptic transmission. *Trends in Neurosciences*, 26(10), 536–542.
- Nieuwenhuys, R. (2001). Neocortical macrocircuits. In G. Roth & M. F. Wullimann (Eds.), *Brain evolution and cognition* (pp. 185–204). New York: Wiley.
- Perea, G., & Araque, A. (2007). Astrocytes potentiate transmitter release at single hippocampal synapses. *Science*, 317, 1083–1086.
- Pezzulo, G. (2008). Coordinating with the future: The anticipatory nature of representation. *Minds and Machines*, 18, 179–225.
- Pezzulo, G., Candidi, M., Dindo, H., & Barca, L. (2013). Action simulation in the human brain: Twelve questions. *New Ideas in Psychology*. <http://dx.doi.org/10.1016/j.newideapsych.2013.01.004>
- Powers, W. T. (1973). *Behavior: The control of perception*. Chicago: Aldine.
- Ransom, B., Behar, T., & Nedergaard, M. (2003). New roles for astrocytes (stars at last). *Trends in Neurosciences*, 26(10), 520–522.
- Roberts, A., & Bush, B. M. H. (Eds.). (1981). *Neurons without impulses*. Cambridge: Cambridge University Press.
- Roesch, E. B., Nasuto, S. J., & Bishop, J. M. (2012). Emotion and anticipation in an enactive framework for cognition (response to Andy Clark). *Frontiers in Psychology*, 3, 1–2.
- Slezak, M., & Pfreger, F. W. (2003). New roles for astrocytes: Regulation of CNS synaptogenesis. *Trends in Neurosciences*, 26(10), 531–535.
- Sokolov, E. M. (1960). Neuronal models and the orienting reflex. In M. Brazier (Ed.), *The central nervous system and behavior* (pp. 187–276). New York: Josiah Macy Jr. Foundation.
- Tani, J., & Nolfi, S. (1999). Learning to perceive the world as articulated: An approach for hierarchical learning in sensory-motor systems. *Neural Networks*, 12, 1131–1141.
- Varela, F. J. (1979). *Principles of biological autonomy*. New York: North Holland.
- Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, 34, 72–87.
- Verkhatsky, A., & Butt, A. (2007). *Glial neurobiology*. Chichester: Wiley.
- Viggiano, D., Ibrahim, M., & Celio, M. R. (2000). Relationship between glia and the perineuronal nets of extracellular matrix in the rat cerebral cortex: Importance for volume transmission in the brain. In L. F. Agnati, K. Fuxe, C. Nicholson, & E. Syková (Eds.), *Volume transmission revisited* (Progress in brain research, Vol. 125, pp. 193–198). Amsterdam: Elsevier.
- Weber, A., & Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1, 97–125.