Charting the Heraclitean Brain: Perspectivism and Simplification in Models of the Motor Cortex

Mazviita Chirimuuta

Abstract

The mainstream theory of the motor cortex is a computational and representational one, but it has been called into question because of an unresolved dispute over whether neurons in this region represent patterns of muscle activation or other movement parameters such as limb velocity. Some neuroscientists have recently argued that the representational theory should be replaced with a dynamical systems one. Here I argue that both of these scientific perspectives are responses to the challenge of constructing relatively simple models of the brain, a system which is extremely complex in the "Heraclitean" sense that its structure and functions are continually changing with experience and never precisely re-occur. Because each of these perspectives employs different assumptions in order to make the task of modeling the brain tractable, they result in different and apparently inconsistent "pictures" of what the brain/mind is like—either an information processor analogous to a computer or a non-representational dynamical system. I discuss proposals to reconcile these perspectives and consider whether any non-perspectival insights about the nature of the brain/mind can be derived from these models.

1. Perspectivism and the Demands of Simplification

It is a frequently stated fact that the human brain is the most complicated object in the known universe; yet it is unclear whether or not this "fact" is a by-product of human vanity or hype drummed up by neuroscientists. To the neurologist Kurt Goldstein it was not obvious that an invertebrate is more simple than a man, and for that reason he saw no obstacle to selecting the human being as his model organism (Goldstein 1939, 2).[1]

The point is that, since everything in nature is in its own way complicated, claims for complexity need to be made specific. Here I will argue that an important respect in which the brain is complex is that it is *Heraclitean*. By this I mean that the brain, like the Heraclitean river, is the kind of thing that can only maintain its identity by undergoing continual change. John Dupré (2012) has argued that all living organisms should be characterized in this way, that is, as processes rather than entities. In addition, Peter Godfrey-Smith (2016) proposes that the Heraclitean nature of biological cells has important implications for how we understand cognition: it is the important difference between the nervous system and man-made computers.

The fact that the brain is made of living tissue such as neurons (an electrically excitable type of biological cell), means that its constitution is constantly changing with metabolism. As Marder and Goaillard (2006, 563) describe, "each neuron is constantly rebuilding itself from its constituent proteins, using all of the molecular and biochemical machinery of the cell." Godfrey-Smith's idea is that the inherent changeability of biological tissue was leveraged during the evolution of the nervous system, as a means for learning and coping with the challenge of staying alive in an unstable world, and is an important factor in making biological "computation" what it is. This, he argues, puts limits on the functional equivalence between brains and artificial computers.

Looking outside the cranium, John Haugeland (1996) has argued that the fact that the brain is densely inter-connected with a mobile body, itself operating in an ever-changing environment, means that the mappings between "inner" neural states and "external" consequences may be

---

[1] In the case of *C. elegans*, a tiny worm, it seems that Goldstein is vindicated by current neuroscience: "A counterintuitive finding in C. elegans is that there is no such thing as 'simplicity' despite the reduced connectome (302 neurons, 6963 synapses, 890 gap junctions), even at the earliest stage of sensory processing." (Frégnac 2017, 473)

constantly evolving. This is more contentious than the biological argument because, as we will see, the empirical evidence for the presence or lack of stable mappings is not conclusive.

What is uncontroversial, however, is that science thrives when complex things can be made to seem simple. Various authors have made the case that complex systems, especially those studied in the biological and behavioral sciences, afford modeling from a variety of perspectives because no one set of theoretical or experimental practices gives the scientist access to all of the relevant phenomena in the domain of interest.[2] Here I build on this work by emphasizing not only the way that scientific perspectives "passively" filter out details not relevant to their own theory and practice but also the way that they "actively" impose simplifying assumptions onto the target system.[3]

My case study in section 2 presents the ongoing controversy between two perspectives on the motor cortex as amounting to a difference in choice of methods selected to cope with the brain's Heraclitean nature. The assumptions embedded in each modeling perspective are effective ways to simplify the brain, but they lead to views about the function of motor cortex that are apparently contradictory. In section 3 I discuss the philosophical implications of this clash of perspectives. Does it follow that there are non-perspectival truths about the motor cortex (and by extension, the primate brain) that are unknowable to science? In section 4 I voice some support for this Kantian conclusion by presenting a general framework for thinking about the operation of abstraction in science.

2. Two Perspectives on the Motor Cortex

[2] See, e.g., Mitchell (2003) and Longino (2013).
[3] Philosophers of science typically refer to the former process as "abstraction" and the latter as "idealization." I think that the distinction between the two is less clear-cut than is normally supposed, so the abstractions I discuss below should not be thought of as *pure* simplifications.

Since everything in nature is complicated in its own way, it is important to recognise simplicity for what it is — something manufactured by means of the scientific method, both materially (in the setting up of laboratory condition) and conceptually (by devising abstract and idealized mathematical representations [Cartwright 1999]). Take the proverbial Heraclitean flux: "on those stepping into rivers staying the same other and other waters flow" (Graham 2015). The scientist dipping her toe in the waters finds, indeed, that in a global sense the river stays the same and there is a regularity to its undulations; yet, the constituents of the river in terms of which she would wish to explain its global properties and regularities are themselves ever changing. This poses a challenge which can be met with various strategies, two of which I will discuss here because they are analogues to the two perspectives on the motor cortex that form my case studies.

The first strategy is to approximate to stasis. One constructs a model of the river at a snapshot of time, ignoring its dynamics and changes in composition. The other strategy is to find simple flow patterns. In this case the dynamics are the target of the model, but rather than attempting to represent every tiny current and eddy, one seeks a compact representation of only the major currents that may be repeatedly observed such as the one parallel to the bank and any due to large features such as islets in the stream.

Likewise, when neuroscientists take on the challenge of modeling the Heraclitean brain, a popular strategy is to assume that the response properties of neurons are approximately fixed—that the neural system is one with stable input-output relationships, which can be represented as a mathematical function. In essence, the proposal is that each neuron represents or codes for some state of affairs in the extra-cranial world. This is the *Intentional Perspective.* An alternative is to model the dynamical evolution of the neural system but seek a relatively simple set of equations governing it. This is the *Dynamical Perspective*, and it often (but not always) comes with the denial that neurons code for or represent anything. The neuroscience of the motor cortex is a particularly apt topic here, because there has been a longstanding and often heated dispute over what the function of this brain area is, leading the scientists themselves to be explicit in stating and arguing for their different theoretical perspectives.[4]

2.1 The Intentional[5] Perspective

Near the start of his lecture arguing that the basis of the success of the exact sciences is their ability to find "economical" representations of phenomena, Ernst Mach (1895, 186) tells us

> 'Life understands not death, nor death life.' … Yet in his unceasing desire to diminish the boundaries of the incomprehensible, man has always been engaged in attempts to understand death by life and life by death.

With these words, Mach foresaw why the information-processing and intentional approach would be such a dominant force in the neuroscience of the future. The point is that even though the living brain is inherently Heraclitean, treating it as functionally equivalent to a non-living symbolic system (a computer) has been an effective way to abstract away from the fluid details of neural hardware.

Some of the founders of computational neuroscience, such as Rashevsky, McCulloch, and Pitts, were quite self-aware about the purpose of this kind of abstraction and idealization (Abraham

[4] Omrani et al. (2017) is a very helpful review of the current approaches to motor cortex. There are more options than the two perspectives I present here, but space does not permit discussion of these. See also Scott (2008).

[5] I use the word "intentional" instead of "representational" to avoid confusion with the scientific representations (models, maps, etc.) that come up in the discussion of scientific perspectives. Shenoy et al. (2013) and Churchland et al. (2012) refer to it as the "representational perspective."

2002).[6] On the other hand, neuroscientists and philosophers under the influence of functionalist theories of mind have had more of a tendency to interpret the brain-computer analogy in a literal way: to treat even rough functional equivalence as an indicator of sameness at a higher level of description, namely, at the level of the coding scheme or algorithm that both systems are said to implement.

The intentional perspective on the motor cortex is comparable to the "fixed filters" model of the visual system.[7] The core assumption is that individual neurons represent or code some parameters relevant to movements in specific body parts. These may be individual muscle activations, sequences of muscle activations, or higher-order parameters such as the velocity of an arm movement. For example, the cosine tuning model of Georgopoulos et al. (1986) treats each motor cortical neuron controlling arm movement as firing maximally at its preferred direction of movement, with firing rate dropping away as a cosine function for non-preferred directions. One of the major difficulties for this perspective has been that neurophysiological recordings have yielded partial evidence for each of these hypotheses — and more — about what the motor cortex codes, leading to a lack of consensus within the intentional camp (Omrani et al. 2017). Within the intentional tradition trial-to-trial variability in neuronal responses is classified as *noise,* rather than as variance to be modeled and explained. This is in part for practical reasons (see section 2.2), and in part because a common assumption is that the neurons' tuning properties are fixed, and so variability in responses is not coding anything.

[6] Rashevsky, for example, says: "Following the fundamental method of physicomathematical sciences, we do not attempt a mathematical description of a concrete cell, in all its complexity. We start with a study of highly idealized systems, which at first may not even have any counterpart in real nature." (quoted in Abraham 2002, 16).

[7] I.e., the idea that individual neurons in the visual system selectively respond to a particular kind of stimulus and that these tunings are stable across time — independent of task and stimulus context. See Chirimuuta and Gold (2009) for discussion.

The assumption that motor cortical neurons code for intended movements has found a practical application in brain computer interface (BCI) technologies that record from this brain region and employ decoding algorithms on the data to derive signals for controlling a robot or cursor. However, it does not follow from the fact that neural data can be decoded in these experiments that the intentional models are realistic or even approximately true of the brain. Certain assumptions made by the decoding algorithms have been shown to be false with respect to neurophysiology of the motor cortex, but during the experiments the brain adapts to biases introduced by the models (Koyama et al. 2010).

2.2 The Dynamical Perspective

The dynamical perspective is the more recent arrival, though its advocates credit Thomas Graham Brown (1882–1965), an associate of Charles Sherrington, with having anticipated their central claim. Speaking of the spinal cord, Brown (1914, 40) writes, “the fundamental activity of this system is the rhythmic.” On the current view we are told that the motor cortex is a “pattern-generation machine” (Kaufman in Omrani et al. 2017, 1835).

One way to summarise the difference between the dynamical and the informational perspective is to say that the relationship of *causation* (between neurons and bodily movements) replaces the intentional relation. While all agree that motor cortical activity is causally upstream of movement, proponents of the dynamical view do not give this an intentional spin (viz., positing that causal interactions between neurons and muscles are merely the medium of information transmission). Instead, they treat the cortex and muscles as coupled oscillatory systems, and ask how the cortex orchestrates its sequence of oscillations (of neural population firing) such that they eventually cause an intended sequence of muscle contractions. A basic intuition here is that the oscillations in populations of cortical neurons, at different frequencies and phases, are analogous to a Fourier basis set of sine waves, with which any irregular waveform can be approximated. Likewise, firing patterns in the motor cortex constitute a basis set which, when appropriately deployed, leads to the execution of the range of bodily movements.

Whereas the empirical support for the intentional perspective comes in the form of single neuron tuning curves for movement parameters, the dynamical view has relied on neural population data, processed to show low dimensional structure. These kinds of data analyses have become common elsewhere in neuroscience with the increase in number of neurons simultaneously recorded; they are the characteristic methods of abstraction within the dynamical perspective.[8] If 100 neurons are recorded during one experimental trial (e.g. an arm reach), the resulting dataset has 100 dimensions (one neuron per dimension). But given the correlations between individual neurons, dimensionality reduction techniques such as principle components analysis (PCA) or factor analysis can typically fit the data into a c.10 dimensional space. The dynamics are represented by plotting the activity of the neural population as a trajectory through a low dimensional state space.

One feature of the dynamical perspective is that single neurons lose their privileged status when neuroscientists set about trying to interpret cortical function. It is a prediction of this approach that the firing patterns of many of the individual neurons will not be interpretable in

---

[8] It bears emphasis that the existence of low-dimensional structure in high-dimensional neural population data is a precondition of the dynamical approach. This is how Cunningham and Yu (2014, 1507) compare the intentional and dynamical perspective in terms of their different strategies for simplifying the brain:

> One of the major pursuits of science is to explain complex phenomena in simple terms. Systems neuroscience is no exception, and decades of research have attempted to find simplicity at the level of individual neurons. Standard analysis procedures include constructing simple parametric tuning curves and response fields, analyzing only a select subset of the recorded neurons, and creating population averages … Recently, studies have begun to embrace single-neuron heterogeneity and seek simplicity at the level of the population as enabled by dimensionality reduction.

Cunningham and Yu provide a clear review of the methods described in this paragraph.

terms of external parameters (Kaufman, as quoted in [Omrani et al. 2017, 1835]). Furthermore, Cunningham and Yu (2014, 1501) make the important point that the shift to simultaneous population recordings makes it possible to investigate the causes of trial-to-trial variability, a component of the data that in single-neuron studies is bracketed as noise and dealt with by averaging across multiple trials. This is unavoidable due to the lack of the statistical power in single neuron data that would be needed to support any conclusions as to the source of variability.

Given that the alliance of population data and dynamical modeling[9] has the potential to embrace and explain (rather than average away) some of the trial-to-trial variability in the brain's responses, it may well seem that this perspective takes us closer to the truth of the Heraclitean brain. Indeed, Haugeland (1996, 123), in his recounting of Hubert Dreyfus's challenge to the entire information-processing framework, tells us that a scenario in which there are no stable mappings between brain states and motor outputs (such as letters typed on a keyboard) would completely undermine the notion of a neural code. The resulting picture of embodied, embedded intelligence is one that has been promoted by some practitioners of dynamical modelling in cognitive science.[10] However, it would be far too quick to argue from the existence of trial-to-trial variability in data recorded during repeated movements to the conclusion that there are no roughly stable mappings between patterns of motor cortex activity and resulting bodily movements, that the "content" of the activity patterns is entirely context dependent. For one thing, there are many neurons for which the mappings are reliable enough so that averaging across trials reveals a preference for a particular direction of movement; this is the core result that undergirds the intentional perspective. Also, the cognitive and body context is not the only source of variability; neurophysiological recording

[9] This is not to imply that DST cannot be applied to single-neuron data, as many studies attest.
[10] E.g.,van Gelder (1995) and Chemero (2009).

techniques are also noisy, and so it is an open question how much trial-to-trial variability is due to behavioral context or due to the recording methods.

Now a defender of Haugeland or Dreyfus may reply that if it were possible to observe the motor cortex during naturalistic movement conditions, where cognitive and bodily context is uncontrolled — where attention wanders freely, and the posture of the rest of the body is not constrained by harnesses (as happens during experiments on arm reaching in monkeys) — then the roughly stable mappings would dissipate and be seen for what they are: an artifact of laboratory conditions. This is an interesting conjecture because if the way that motor cortex activity maps to movement in naturalistic conditions is genuinely Heraclitean, while stability is generated by the constraints of laboratory conditions, then it turns out that the dynamical perspective is not much better placed to represent the Heraclitean motor cortex, in all of its changeable glory, than the intentional one.

The reasons are as follows. Firstly, one aim of the population analyses performed by the dynamical camp is to identify reliable correlations between population (as opposed to single neuron) activity and movements. Even though these correlations are interpreted causally rather than intentionally, and such mappings do not require that *individual* neurons behave in the same way on each trial, so long as some global pattern of activity is maintained (e.g., a certain number of neurons oscillating in a particular way), the research program would be a non-starter if there were absolutely no consistent relationships between movements and neural activity, at any level of description. (This is perfectly consistent with the point made above, that given the statistical power afforded by multi-neuron simultaneous recording, this approach makes it possible to explain some trial-to-trial variation in terms of behavioral or cognitive context.)

Secondly, just as the core findings of the intentional perspective (i.e., consistent neuron-movement mappings) may be dependent on the fact that the neural responses are generated in controlled laboratory conditions, the core findings of the dynamical perspective, that the population data can be represented in a low-dimensional state space, which yield hypotheses about the relationship between global state and movement, may themselves be dependent on the same simplifications introduced in the laboratory. Surya Ganguli and colleagues have presented some formal results which, they argue, show that the low-dimensional structure revealed in neural population studies so far is due to the simplicity of tasks used in experiments ([Gao and Gangulli 2015; Gao et al. 2017]; but see [Golub et al. 2018]). The upshot is that even if the brain outside of the laboratory is truly Heraclitean in the way that Haugeland and Dreyfus propose, the very techniques used by neuroscientists, in order to observe its workings, tend to make its behavior less complex than this. The Heraclitean brain, one might say, is not an observable object of science.

3. Relating the Perspectives

Table 1 summarizes the main points of difference between the two perspectives. These concern kinds of experimental protocol and data analysis, as well as the original sources (beyond neuroscience) of the theoretical frameworks. The final line of the table states the two divergent claims about the function of the motor cortex that are made on the basis of these avenues of research: either that this brain area is specialized for coding movements or that it is a pattern generator.

[INSERT TABLE 1.1 HERE]

One way to understand the lesson of perspectivism is that the picture of the world offered by science is like from a cubist painting: a handful of different points of view are employed at once, and reality comes through to us only as fragmented and distorted. We have no access to a God's eye view of nature, and science is not a clear, accurate depiction of reality. One might conclude, therefore, that from the dynamical perspective the motor cortex is a pattern generator, and from the intentional one it is a system for coding movements, and that is the end of the story.[11]

However, many would be dissatisfied if matters were left just there. There are unanswered questions about the relationship between the perspectives: are they competitors or complementary to one another? Could one perspective ultimately subsume the other? Furthermore, an objection to the perspectivist conclusion, as just stated, is that it is just a version of relativism. Thus, it seems, we are owed a more rigorous notion of perspectival truth in order to avoid this outcome. In section 3.2 I will bring some options to the table here. Before that I will examine the relationship between the two perspectives.

3.1 Rivals or Allies?

A theme of Sandra Mitchell's (2003) account of integrative pluralism in science has been that different perspectives are often complementary to one another, with one view compensating for the deficiencies of another, and that cooperation across perspectives occurs when there is a practical challenge that cannot be addressed with one approach alone. Integration does not entail subsumption of one perspective by another; the various models, methods, and representations that constitute a perspective will retain their distinct identities.[12] In keeping with the claims of

[a] See Beer and Williams (2015: 5). They present a modified version of information theory which encompasses such temporal changes.
[b] NB in themselves the data only indicate that there are correlations between neuronal activity and movements. These relationships are then interpreted as intentional or causal ones.
[c] Fairhall (2014: x)
[11] Cf. Giere (2006, 5–6): "the strongest claims a scientist can legitimately make are of a qualified, conditional form: 'According to this highly confirmed theory (or reliable instrument), the world seems to be roughly such and such'."
[12] See (Mitchell 2003, this volume).

integrative pluralism, neuroscientist Adrienne Fairhall has argued that an integration between dynamical and informational approaches is a crucial step in the progress of neuroscience:

> Ultimately, the development of methods to map the dynamics of the physical substrate onto the computational is the bottleneck in our ability to truly comprehend the biological mechanisms of intelligence. (Fairhall 2014, xi)

A condition of integrative pluralism is that there are pros and cons associated with each perspective, and for that reason they mutually support one another. This raises the question of what the strengths and weaknesses of each of the motor cortex perspectives are. I argued above that the dynamical approach cannot be claimed to give us the unvarnished truth about the Heraclitean brain because of its reliance on finding low-dimensional structures. That said, it is more faithful to the ever-flowing nature of brain processes than the intentional perspective, with its static way of conceptualizing neural functions. So why retain the intentional perspective now that the dynamical one has come of age?

There is something important about the brain that the intentional perspective captures but which eludes the dynamical one. This is the fact that brain states really do seem to be directed to (not merely correlated with or caused by) external states of affairs. These states persist in memory, absent the external stimulus, and are robust to perturbations in the brain. For instance, BCI experiments have shown that if the mappings between neural activity and motor output are perturbed, the motor cortex will adjust its activity patterns to achieve a new set of stable mappings (Jarosiewicz et al. 2008). There is no obvious way to describe the apparent purposefulness of this reorganization within the dynamical framework. More generally, directed relationships between neural states, past events, goals, and future expectations do seem to be critical to explaining what makes a system intelligent, hence they have often been taken to be a "mark of the cognitive."

The dynamical perspectives takes the same correlations between neural activity and external states of affairs that undergird the intentional account but lends them no more than causal significance. So it is open to a proponent of the intentional framework to *label* the neural population patterns found using the dynamical methods as *representations of movements*.[13] In the motor cortex case there is nothing to prevent an intentionalist doing this, but there are not the grounds to insist on this re-labeling either.

In the debates that have gone on between proponents of the two perspectives on the motor cortex the tone has not been particularly conciliatory. The background assumption, it seems, is that the two approaches are natural rivals and not allies — that ultimately one view must be right and the other wrong. In particular, advocates of each side have taken pains to show that the core phenomena stated by their opponents are recoverable within their own modeling framework (for example Kaufman in [Omrani et al. 2017]). The title of a comparison study by Michaels et al. (2016) is telling: "Neural population dynamics during reaching are better explained by a dynamical system than representational tuning."

In contrast, one prominent figure in the field of dynamical modelling, Randall Beer, has made the case for an alliance between dynamical and informational approaches. Beer and Williams (2014) show that the behavior of one extremely simple, simulated cognitive agent, can be explained either using the formalism of information theory (IT) or dynamical systems theory (DST). They write:

[13] I thank David Bain, Fiona MacPherson and Scott Sturgeon for pressing this point.

> As mathematical theories, IT and DST can be applied to any system that takes the proper form to meet their defining requirements; they intrinsically make no scientific claim as to 'what's really going on.' Instead, they are best viewed as distinct mathematical lenses through which we can examine the operation of a system of interest. (Beer and Williams 2014, 2)

These authors go on to say that "the mathematical languages themselves are merely more or less useful to a given purpose" (Beer and Williams 2014, 23). Inspired by these remarks, one might attempt a reconciliation between the two perspectives via an instrumentalism which denies that mathematical models afford any view of underlying nature beyond the empirical predictions. As such, the use of these mathematical tools tells us nothing about what the brain is really like.

3.2 On Perspectival Truth

Embracing this conclusion, for the perspectives on the motor cortex, would invite the objection that perspectivism is nothing more than instrumentalism rebranded (Morrison 2011). Michela Massimi is one philosopher of science who has taken pains to show that perspectivism is actually a version of scientific realism, by developing a substantial and non-relativistic notion of perspectival truth. In this section I will examine whether or not her account is applicable to the motor cortex example.

One version of perspectival truth that Massimi (2018, 349) considers is a contextualist one:

> (P_3) *Perspective-dependence$_3$.* Knowledge claims in science are perspective-dependent$_3$ when their truth-conditions depend on the scientific perspective in which such claims are made.

Here, scientific perspectives provide the context in which truth-conditions are defined for the knowledge claims of science. For example, on this account

> <M1 neurons code movement parameters>

would be true in the context of the intentional/informational perspective and false in the dynamical one. The stated benefits of P_3 are that it upholds the realist intuition that science gets things (partially) right but at the same time rejects a monistic view of scientific knowledge in favor of plurality of perspectives that offer "idealized, inaccurate, and yet still true perspectival images of an independent world" (Massimi 2018, 353). The downside, Massimi contends, is that if P_3 is all we aim for, we must also concede that nature-in-itself is an unknowable, noumenal reality. Massimi's more ambitious notion of perspectival truth employs a distinction between context-of-use and context-of-assessment:

> (P_4) *Perspective-dependence$_4$.* Knowledge claims in science are perspective-dependent$_4$ when their truth-conditions … depend on the scientific perspective in which such claims are made. Yet such knowledge claims must also be assessable from the point of view of other … scientific perspectives. (Massimi 2018, 254)

This idea is fleshed out with the example of the claim

<Water is a liquid with viscosity>

which is true from the perspective of hydrodynamics but false according to statistical mechanics. However, if statistical mechanics is deployed as a "context of assessment" for hydrodynamics it can be shown that the property of viscosity is still recoverable in statistical mechanics "as a derivative property" (Massimi 2018, 354). Thus, the knowledge claim of one perspective is validated by the other perspective after all.

There is a parallel to Massimi's account of viscosity in the motor cortex case: proponents of each perspective do claim to be able to recover the core phenomena of the alternative perspective using their own models and assumptions. That is, advocates of the dynamical view have emphasized that the correlations between movements and neural activity, which are taken by intentionalists to be the signature of coding, are also predicted by the dynamical account. Likewise, neuroscientists defending the intentionalist framework hastened to show that curved trajectories in the low dimensional jPC space, argued by Churchland et al. 2012 to be evidence for the dynamical view, were consistent with cosine tuning in the motor cortex (see Michaels et al. 2016). Yet the conclusion that each side draws from any instances of cross validation between perspectives is not the conciliatory or pluralist one; rather such findings are claimed to *undercut* the other perspective. The logic is that if an alternative perspective is not needed to explain a portion of the observed findings, then only one perspective should be employed.

I should emphasize that there may well be sociological and psychological reasons why the debate over motor cortex perspectives has more of the look of a turf war than a display of polite recognition of the need for diversity of theories and methods in science. But one philosophical explanation for why diversity is not met with tolerance here is that there is no big picture of neural and cognitive function that the different perspectives are converging on. From each individual perspective you get an "interpretable" picture of what the motor cortex fundamentally is, but when placed together, with the pluralist claim that both are in some sense true, the picture becomes incoherent.

The situation is illustrated in figure 1. At the level of mathematical formalism, the intentional and DST perspectives do cross-validate one another in the way that Massimi requires for her robust notion of perspectival truth. As Beer and Williams (2014) show for their minimal cognitive agent, some of the phenomena isolated by dynamical modelling also show up in the informational model and vice versa. However, when one moves beyond the pure formalism, to the level of interpretation of the models,[14] the perspectives diverge. Figure 1 presents additional layers of interpretation which end ultimately with different philosophical views about what the brain/mind fundamentally is.

[INSERT FIGURE 1.2]

Figure 1. Illustration of cross-validation of quantitative models and divergence of qualitative interpretations of those models. Beer and Williams (2014) provide a demonstration of cross-perspective validation for their very simple, minimally cognitive agent.

Each perspective provides mathematical formalisms for describing neural activity. At this purely quantitative level, the perspectives can be shown to be consistent with one another, satisfying P_4.

[14] By "interpretation" here I mean, roughly, what is implied by the model over and above its quantitative predictions; cf. the semantic content that scientific realists — but not instrumentalists — would associate with scientific theories. Interpretation can be thought of as an informal representation of the system. Note that for the scientific realist — but not the instrumentalist — the mathematical formalism can also be taken to provide a representation of the neural system.

When one considers the qualitative descriptions of the neural systems associated with each perspective (the claims about what the cortex "is like"), over and above the mathematical formalism, inconsistency appears. Note that *functionalism* is a philosophical theory of the mind-body relationship which often presupposes internal representations; in contrast, the *embodied-embedded cognition* theory denies that representations are needed in order to explain mental capacities. I do not suppose that many neuroscientists employing these models also commit themselves to these philosophical theories, but it is relevant that the models do lend themselves to these higher levels of interpretation.

Since the clash of perspectives only comes with the interpretation of the mathematical models, one response to this difficulty is to strongly discourage interpretation. The task of the neuroscientist is to shut up and calculate, and leave speculation about the nature of the brain/mind to the philosophers. This response is in keeping with Beer and William's recommendation that the mathematical theories of IT and DST by themselves make no claims about "what's really going on."

A pluralism that blends into instrumentalism by putting restrictions on extra-empirical interpretation is not an option for Massimi, since a non-negotiable claim of scientific realism is that the successful theories and models employed by scientists can also be interpreted in order to tell us something about the underlying nature of things. Massimi argues that for perspectivism to be made compatible with realism, the different perspectives must be shown to endorse each other's knowledge claims; but in our case the mutual reinforcement is only possible at the level of uninterpreted mathematical formalism. Once interpretation is lent to the models, the clash of perspectives is jarring.

This amounts to a dilemma for the ambitious perspectival realism advocated by Massimi. P_4 requires the knowledge claims of each perspective to be endorsed by the other perspectives. In the IT/DST example the endorsement is granted only if the cross-perspective assessment is restricted to the quantitative predictions of the mathematical models, while no cross-perspective endorsement is granted to the qualitative interpretations associated with those models — the bigger picture they offer as to "what's really going on" in the motor cortex, and brain more generally. So *either* P_4 knowledge claims must be restricted to the quantitative predictions of the mathematical models, leading to instrumentalism or, if not so restricted, it is clear that convergence of knowledge claims across perspectives does not obtain; this implies that one of the perspectives is the truer one and pluralism must be abandoned.[15] Thus, it seems, the ambitious notion of perspectival truth, P_4, is not applicable to our case. I will conclude the chapter with some thoughts on why the more modest P_3 is not such a bad notion to settle for.

4. Conclusion. Two Philosophical Perspectives on Abstraction

It is interesting that the word "abstraction" bears two different meanings in contemporary philosophy: one lofty, the other mundane. In the lofty sense, an abstraction is an abstract entity, not spatially and temporally located and so possibly residing in Plato's heaven. Speaking mundanely (the use more common amongst philosophers of science), abstraction is synonymous with simplification and paired with idealization, an important model-building strategy employed by scientists. These two conceptions of abstraction animate two very different explanations for the "unreasonable effectiveness of mathematics"; that is, they provide two different answers to the question of why mathematics is such a useful tool in science. The lofty explanation is that the

[15] That is not to imply that anyone knows which perspective is the truer one. Furthermore, it could also be that all of the current perspectives are equally far from the truth.

underlying reality of nature consists in mathematical structure, and the task of the exact sciences is to discover these. The mundane one is that science progresses when humans find ingenious ways to simplify complex phenomena, and mathematics is the pre-eminent tool for doing this. Adherents to the lofty way of thinking about abstraction and the role of mathematics in science are in good company — not only Plato but Galileo (stating that the book of nature is written in the language of mathematics), Descartes (and other rationalist philosophers), and contemporary ontic structural realists have intellectual kinship here. However, this Platonic tradition meets difficulty with the existence of pluralities of different kinds of mathematical representation of natural phenomena. If the construction of a predictively powerful model of natural phenomena is also, in some sense, a revelation of the mathematical laws that underlie observable phenomena, then how can it be that the book of nature seems to be written by multiple authors?[16]

Kant is a figurehead for the mundane approach. Instead of taking abstract mathematical representation to be somehow more "true" than the concrete, observable phenomena, on this account one regards the mathematical representations as a set of structures employed by the human mind as a means to order the observable phenomena.[17] When confronted with the varying, multifaceted and disordered phenomena in nature, mathematics offers a useful set of structures for

[16] Given Beer and Williams' finding of the compatibility of the two formalisms, someone sympathetic to the Platonic view might respond here that, yes, the brain partakes of these two different formalisms and we should desist from trying to interpret them and reconcile the interpretations. My response would again be that imposing any such ban on interpretation leads you into instrumentalism. As Robert Briscoe and Will Davies have pointed out to me, more problematic cases for my arguments are ones where scientists have endeavoured to merge the perspectives both in terms of formalism and interpretation. In the case of dynamical and intentional models of the motor cortex, I have not yet come across work of this kind.

[17] NB my characterization of a Kantian view is not aiming at accurate exegesis of Kant's thought on mathematical science, nor does it come with any commitment to Kantianism regarding the ontology of numbers or epistemology of mathematics. For instance, one could think that mathematics is learned by the mind's apprehension of Platonic Forms but still be a "Kantian" in the sense relevant here, i.e., by denying that mathematical structures are the truer reality underlying the appearances in nature and asserting that the utility of maths in science comes from the mind's ability to employ certain simple structures in its apprehension of nature. However, there is a connection between the Platonic tradition I characterize here and Platonism regarding the ontology of numbers, in that the indispensability arguments for the existence of numbers presuppose the lofty explanation for the success of mathematical science. I thank Alastair Wilson for this point.

imposing representational order on them, especially by means of leaving out details — the process of abstraction. Once one makes the Kantian move of looking "inwards" for the explanation of how mathematical representations yield knowledge of nature it is not jarring or surprising that there are multiple ways to achieve order and abstract, hence there may be a plurality of kinds of mathematical model for the same piece of nature.

This way of thinking about abstraction and the scientific method was popular in the philosophy of science a century ago, even if neglected now.[18] I believe it offers significant benefits for thinking about perspectival pluralism. Not only does it welcome the existence of multiple perspectives, it also permits a relaxed response to the possibility of there being unknowable, non-perspectival truths. In our case we can say that the brain-in-itself is not knowable in its full, Heraclitean complexity because no human scientist (with limited, human cognitive powers) would be able to theorise it completely and accurately as such. The brain-in-itself is not mysteriously unthinkable; it is just very complicated. At the same time, the Heraclitean brain provides constraints on what counts as an acceptable representation of it, and this means that contextual knowledge claims about it — as in P_3 — are not fictitious or relativistic ones. To see as through a glass darkly is still to see something.[19]

[18] I am thinking here of Duhem (1954), Cassirer (1957), Husserl (1970), and Whitehead (1938). In particular, the pluralism advocated here is more in the spirit of neo-Kantians such as Cassirer than Kant himself. Cassirer (1957, 409) makes a point highly relevant to my study, that the exact sciences are only concerned with events "under the aspect of [their] repeatability." We can say that the Heraclitean brain in its never-exactly-repeating richness is simply not a concern to the mathematical neuroscientist; the brain must be presumed, qua object of mathematical neuroscience, to be non-Heraclitean.

[19] We might also think of these as two theological outlooks. On the Platonic side, God is conceived of as a mathematically informed creator, and the structure of his creation is revealed through functions and number. Creation is intelligible to human reason to extent that we understand mathematics. On the other we have a negative theology, where God and his ways are somewhat inscrutable to the finite human mind:

> "He [the scientist] will choose a certain formula because it is simpler than the others; the weakness of our minds constrains us to attach great importance to considerations of this sort. There was a time when physicists supposed the intelligence of the Creator to be tainted with the same debility." (Duhem 1954, 171)

Acknowledgements

I would like to thank audience members at the Philosophy, Psychology and Neuroscience Research Seminar (University of Glasgow), Philosophy Work in Progress Seminar (University of Birmingham), the Perspectivism Reading Group (University of Pittsburgh), and the Neural Mechanisms Online (University of Turin) seminar series for lively discussions and many useful suggestions. I am also grateful to Mark Churchland for a discussion of his research, and to the editors of this volume for their many helpful comments on the text.

References

Abraham, T. H. 2002. "(Physio)logical Circuits: the intellectual origins of the McCulloch-Pitts Neural Networks." *Journal of the History of the Behavioral Sciences* 38(1): 3–25.

Beer, R. D., and Williams, P. L. 2015. "Information processing and dynamics in minimally cognitive agents." *Cognitive Science* 39(1): 1–38.

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." *Journal of Physiology* 48(1): 18–46.

Cartwright, N. D. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Cassirer, E. 1957. *The Philosophy of Symbolic Forms; Volume 3: The Phenomenology of Knowledge*. New Haven: Yale University Press.

Chemero, A. 2009. *Radical Embodied Cognitive Science* Cambridge: MIT Press.

Chirimuuta, M., and Gold, I. J. 2009. "The embedded neuron, the enactive field?" In *The Oxford Handbook of Philosophy and Neuroscience*, edited by Bickle, J., 200–225. Oxford: Oxford University Press.

Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., and Shenoy., K. V. 2012. "Neural population dynamics during reaching." *Nature* 487: 51–56.

Cunningham, J. P., and Yu, B. M. 2014. "Dimensionality reduction for large-scale neural recordings." *Nature Neuroscience* 17: 1500–1509.

Duhem, P. 1954. *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press.

Dupré, J. 2012. *Processes of Life* Oxford: Oxford University Press.

Fairhall, A. 2014. "The receptive field is dead. Long live the receptive field?" *Current Opinion in Neurobiology* 25: ix–xii.

Frégnac, Y. 2017. "Big data and the industrialization of neuroscience: A safe roadmap for understanding the brain?" *Science* 358(6362): 470–477.

Gao, P., Trautmann, E., Yu, B. M., Santhanam, G., Ryu, S. I., Shenoy, K. V., and Ganguli, S. 2017. "A theory of multineuronal dimensionality, dynamics and measurement." bioRxiv 214262. https://doi.org/10.1101/214262.

Georgopoulos, A., Schwartz, A., and Kettner, R. 1986. "Neuronal population coding of movement direction." *Science* 233(4771): 1416–1419.

Giere, R. 2006. "Perspectival pluralism." In *Scientific Pluralism*, edited by Kellert, S. H., Longino, H. E., and Waters, C. K., 26–41. Minneapolis: University of Minnesota Press.

Godfrey-Smith, P. 2016. "Mind, Matter, and Metabolism." *Journal of Philosophy* 63:481-506.

Goldstein, K. 1939. *The Organism: a holistic approach to biology derived from pathological data in man*. Salt Lake City: American Book Company.

Golub, M. D., Sadtler, P. T., Oby, E. R., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., Batista, A. P., Chase, S. M., and Yu, B. M. 2018. "Learning by neural reassociation." *Nature Neuroscience* 21: 607–616.

Graham, D. W. 2015. "Heraclitus." *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2015/entries/heraclitus/>

Haugeland, J. 1996. "Body and world: a review of *What Computers Still Can't Do: A Critique of Artificial Reason* (Hubert L. Dreyfus)." *Artificial Intelligence* 80: 119–128.

Husserl, E. 1970. *The Crisis of European Sciences and Transcendental Phenomenology*. Evanston: Northwestern University Press.

Jarosiewicz, B., Chasea, S. M., Fraser, G. W., Velliste, M., Kass, R. E., and Schwartz, A. B. 2008. "Functional network reorganization during learning in a brain-computer interface paradigm." *Proceedings of the National Academy of Sciences* 105(49): 19486–19491.

Koyama, S., Chase, S. M., Whitford, A. S., Velliste, M., Schwartz, A. B., and Kass, R. E. 2010. "Comparison of brain-computer interface decoding algorithms in open-loop and closed-loop control." *Journal of Computational Neuroscience* 29(1–2): 73–87.

Longino, H. E. 2013. *Studying Human Behavior*. Chicago: Chicago University Press.

Mach, E. 1895. "On the Economical Nature of Physical Enquiry." In *Popular Scientific Lectures*, translated by McCormack, T. J., 186–213. Chicago: Open Court.

Marder, E., and Goaillard, J. M. 2006. "Variability, compensation and homeostasis in neuron and network function." *Nature Reviews Neuroscience*, 7: 563–574.

Massimi, M. 2018. "Four Kinds of Perspectival Truth." *Philosophy and Phenomenological Research* 96(2): 342–359.

Michaels, J. A., Dann, B., and Scherberger, H. 2016. "Neural Population Dynamics during Reaching Are Better Explained by a Dynamical System than Representational Tuning." *PLOS Computational Biology* 12(11): e1005175.

Mitchell, S. D. 2003. *Biological Complexity and Integrative Pluralism*. Cambridge: Cambridge University Press.

Morrison, M. 2011. "One phenomenon, many models: Inconsistency and complementarity." *Studies in History and Philosophy of Science Part A* 42(2): 342–351.

Omrani, M., Kaufman, M. T., Hatsopoulos, N. G., and Cheney, P. D. 2017. "Perspectives on classical controversies about the motor cortex." *Journal of Neurophysiology* 118(3): 1828–1848.

Scott, S. H. 2008. "Inconvenient Truths about neural processing in primary motor cortex." *The Journal of Physiology* 586(5): 1217–1224.

Shenoy, K. V., Sahani, M., and Churchland, M. M. 2013. "Cortical control of arm movements: a dynamical systems perspective." *Annual Review of Neuroscience* 36: 337–359.

van Gelder, T. 1995. "What might cognition be if not computation?" *The Journal of Philosophy* 92(7): 345–381.

Whitehead, A. N. 1938. *Science and the Modern World*. New York: MacMillan.