

# The Philosophy and Neuroscience Movement

**Morteza Alibakhshi Kenari**

Martyr Beheshti University of Medical Sciences and Health Services, Tehran, Iran

Email: [Morteza.Alibakhshikenari@gmail.com](mailto:Morteza.Alibakhshikenari@gmail.com)

Received 22 March 2014; revised 28 April 2014; accepted 6 May 2014

Copyright © 2014 by author and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

A movement dedicated to applying neuroscience to traditional philosophical problems and using philosophical methods to illuminate issues in neuroscience began about thirty-five years ago. Results in neuroscience have affected how we see traditional areas of philosophical concern such as perception, belief-formation, and consciousness. There is an interesting interaction between some of the distinctive features of neuroscience and important general issues in the philosophy of science. And recent neuroscience has thrown up a few conceptual issues that philosophers are perhaps best trained to deal with. After sketching the history of the movement, we explore the relationships between neuroscience and philosophy and introduce some of the specific issues that have arisen.

## Keywords

Philosophy, Neuroscience, Movement, Consciousness, Science, Relationship, Interaction, Illuminate Issues

---

## 1. Introduction

The exponentially-growing body of work on the human brain of the past few decades has not only taught us a lot about how the brain does cognition, but also had a profound influence on other disciplines that study cognition and behavior. A notable example, interestingly enough, is philosophy. A small movement dedicated to applying neuroscience to traditional philosophical problems and using philosophical methods to illuminate issues in neuroscience began about twenty-five years ago and has been gaining momentum ever since. The central thought behind it is that certain basic questions about human cognition, questions that have been studied in many cases for millennia, will be answered only by a philosophically sophisticated grasp of what contemporary neuroscience is teaching us about how the human brain processes information (Kan, 2011).

The evidence for this proposition is now overwhelming. The philosophical problem of perception has been transformed by new knowledge about the vision systems in the brain. Our understanding of memory has been deepened by knowing that two quite different systems in the brain are involved in short- and long-term memory. Knowing something about how language is implemented in the brain has transformed our understanding of the structure of language, especially the structure of many breakdowns in language, and so on. On the other hand, a great deal is still unclear about the implications of this new knowledge of the brain. Are cognitive functions localized in the brain in the way assumed by most recent work on brain imaging? Does it even make sense to think of cognitive activity being localized in such a way? Does knowing about the areas active in the brain when we are conscious of something hold any promise for helping with long-standing puzzles about the nature and role of consciousness ([Kazu, 2011](#))?

As a result of this interest, a group of philosophers and neuroscientists dedicated to informing each other's work has grown up. Many of these people now have Ph.D. level training or the equivalent in both neuroscience and philosophy.

Much of the work that has appeared has been clustered around five themes.

- Data and theory in neuroscience;
- Neural representation and computation;
- Visuomotor transformation;
- Colour vision;
- Consciousness.

And two big issues lie in the substructure of all of them,

- The relationship of neuroscience to the philosophy of science;
- And whether cognitive science will be reduced to neuroscience or eliminated by it ([Luck, 2010](#)).

We will take up these themes shortly.

## 2. Neuroscience and the Philosophy of Science

In much early philosophy of science, the notion of law is central, as in the Deductive-Nomological theory of scientific explanation or the Hypothetico-Deductive theory of scientific theory development or discussions of intertheoretic reduction. While the nomological view of science seems entirely applicable to sciences such as physics, there is a real question as to whether it is appropriate for life sciences such as biology and neuroscience. One challenge is based on the seeming teleological character of biological systems. Aldrich argue that a teleological approach can integrate neuroscience, psychology and biology ([Aldrich, 2006](#)).

Another challenge to the hegemony of nomological explanation comes from philosophers of neuroscience who argue that explanations in terms of laws at the very least need to be supplemented by explanations in terms of mechanisms ([Almekhlafi, 2004](#)). Here is how their story goes. Nomological explanations, as conceived by the Deductive-Nomological model, involve showing that a description of the target phenomenon is logically deducible from a statement of general law. Advocates of the mechanistic model of explanation claim that adequate explanations of certain target phenomena can be given by describing how the phenomena results from various processes and sub-processes. For example, cellular respiration is explained by appeal to various chemical reactions and the areas in the cell where these reactions take place. Laws are not completely abandoned but they are supplemented ([Barta, 2002](#)).

One main reason why neuroscience raises issues such as these in stark form is that, while there is clearly an enormous amount of structure in the brain (the human brain is made up of roughly 100,000,000,000 neurons), neuroscience has had very little success in finding general laws that all or nearly all brains instantiate. Maybe for at least the complex kinds of activity that underpin cognition, it will turn out that there are no such laws, just masses and masses of individually-distinct (though still highly structured) events.

A related challenge to logical positivist philosophy of science questions whether scientific theories are best considered to be sets of sentences, for example, suggests that the vector space model of neural representation should replace the view of representations as sentences ([Birch, 2008](#)). This would completely recast our view of the enterprise of scientific theorizing, hypothesis testing, and explanation. This challenge is directly connected to the next issue.

## 3. Reduction versus Elimination

There are three general views concerning the relation between the psychological states posited by cognitive sci-

ence and the neurophysiological processes studied in the neurosciences:

1) *The autonomy thesis*: While every psychological state may be (be implemented by, be supervenient on) a brain state, types of psychological states will never be mapped onto types of brain states. Thus, each domain needs to be investigated by distinct means, cognitive science for cognitively-delineated types of activity, neuroscience for activities described in terms of brain processes and structures (Fodor, 2010).

Analogy: every occurrence of red is a shape of some kind, but the colour-type, redness, does not map onto any shape-type. Colours can come in all shapes and shapes can be any colour (Bonk, 2000).

2) *Reductionism*: The types of psychological states will ultimately be found to be a neurophysiological states; every cognitively-delineated type can be mapped onto some type of brain process and/or structure with nothing much left over. The history of science has been in no small part a history of such reduction, as they are (somewhat misleadingly) called (misleading because the reduced kinds still continue to exist): Chemistry has been shown to be a branch of physics, large parts of biology have been shown to be a branch of chemistry, and so on. The reductivists about cognition (or psychology generally) believe that cognition (and psychology generally) will turn out to be a branch of biology.

3) *Eliminativism (aka eliminative materialism)*: Psychological theories are so riddled with error and psychological concepts are so weak when it comes to building a science out of them that psychological states are best regarded as talking about nothing that actually exists.

In the space we have, we cannot go into the merits of reductivist vs. eliminativist claims, but notice that the truth of eliminativism will rest on at least two things:

1) The first concerns what the current candidates for elimination actually turn out to be like when we understand them better. For example, eliminativists about folk psychology often assume that folk psychology views representations as structured something like sentences and computations over representations to very similar to logical inference (Bowman, 2008). Now, there are explicit theories that representation is like that. But it is not clear that any notion of what representations are like is built into our very folk concept of representation. The picture of representation and computation held by most neuroscientists is very different from the notion that representations are structured like sentences, as we will see when we get to computation and representation, so if the sententialist idea were built into folk psychology, then folk psychology would probably be in trouble. But it is not clear that any such idea is built into folk psychology.

2) The second thing on which the truth of eliminativism will depend is what exactly reduction is like. This is a matter of some controversy (Cassell, 2006). For example, It can reductions be less than smooth, with some bits reduced, some bits eliminated, and still count as reductions? Or what if the theory to be reduced must first undergo some rigging before it can be reduced? Can we expect theories dealing with units of very different size and complexity (as in representations in cognitive science, neurons in neuroscience) to be reduced to one another at all? And how much revision is tolerable before reduction fails and we have outright elimination and replacement on our hands (Bickle, 2012)? Deane argues that reductions are usually between theories at roughly the same level (intratheoretic), not between theories dealing with radically different basic units (Deane, 2001).

## 4. Data and Theory in Neuroscience

In a variety of ways, the advent of sophisticated imaging of brain activity has created a new reliance on introspection—it is difficult if not impossible to relate what is going on cognitively to various brain activities without self-reports from subjects about what are going on in them. Introspection has been in bad as a research tool for over 100 years. It has variously been claimed that:

- 1) Introspective claims are unreliable because they are not regularly replicated in others.
- 2) Subjects confabulate (make up stories) about what are going on in themselves when they need to do so to make sense of behaviour.
- 3) Introspection has access only to a tiny fraction of what is going on in oneself cognitively.
- 4) It is impossible for introspection to access brain states.

## 5. Neural Representation and Computation

Neural representation and computation is a huge topic, as we said. We will start with neural representation. The neurophilosophical questions concerning computation and representation nearly all assume a definition of computation in terms of transformation of representations. Thus, most questions concerning computation and repre-

sentation are really questions about representation. Contributions to this topic can be thought of as falling into three groups, though the boundaries between them are far from crisp. There are questions to do with architecture, question to do with syntax, and questions to do with semantics. The question of architecture is the question of how a neural system having syntax and semantics might be structured. The question of syntax is the question of what the formats or permissible formats of the representations in such a system might be and how representations interact with each other based on their form alone. The questions of semantics is the question of how it is that such representations come to represent—how they come to have content, meaning.

### 5.1. Architecture of Neural Representation

Here is some of the thinking afoot currently about neural architecture. Past approaches to understanding the mind, including: Symbolicism, connectionism, Dynamicism. A much less metaphorical approach, or so it is claimed, unifies representational and dynamical descriptions of the mind. First, representation is rigorously defined by encoding and decoding relations. Then, the variables identified at higher levels are treated as state variables in control theoretical descriptions of neural dynamics. Given the generality of control theory and representation so defined, it is claimed that this approach is sufficiently powerful to unify descriptions of cognitive systems from the neural to the psychological levels. If so, contrary could have both representation and dynamics in cognitive science (van Gelder, 2009).

### 5.2. Neural Syntax

The standard way of interpreting synaptic events and neural activity patterns as representations is to see them as constituting points and trajectories in vector spaces. The computations that operate on these representations will be show as vector change (Funk, 2008). This is thus the view adopted in much neural network modelling (connectionism, parallel distributed processing). The system is construed as having network nodes (neurons) as its basic elements and representations are states of activations in sets of one or more neurons (Elwood, 2004).

Recently, work in neural modelling has started to become even more finegrained. This new work does not treat the neuron as the basic computational unit, but instead models activity in and interactions between patches of the neuron's membrane (Goos, 2007). Thus, not only are networks of neurons viewed as performing vector transformations, but so are individual neurons. Neural syntax consists of the study of the information-processing relationships among neural units, whatever one takes the relevant unit to be. Any worked-out story about the architecture of neural representation will hold implications for neural syntax, for what kind of relationships neural representations will have to other neural representations such that they can be combined and transformed computationally.

### 5.3. Neural Semantics

Proponents of functional role semantics propose that the content of a representation, what it is about, is determined by the functional/causal relations it enters into with other representations (Granič, 2007). For informational approaches, a representation has content, is about something, in virtue of certain kinds of causal interactions with what it represents (Dretske, 2010). In philosophy of neuroscience, Funk has subscribed to functional role semantics at least since 2011. His account is further fleshed out in terms of state-space semantics (Funk, 2011).

The neurobiological paradigm for informational semantics is the feature detector, for example, the device in a frog that allows it to detect flies (Hayes, 2009). Establishing that something has the function of detecting something is difficult. Mere covariation is often insufficient. Jumani identified receptive fields of neurons in striate cortex that are sensitive to edges (Jumani, 2010). Did they discover edge detectors? Challenge the idea that they had, showing that neurons with similar receptive fields emerge in connectionist models of shape-from-shading networks (Lehky, 2011). Akins offers a different challenge to informational semantics and the feature detection view of sensory function through a careful analysis of therm operation. She argues that such systems are not representational at all (Akins, 2011).

As was true of neural syntax, any worked-out story about the architecture of neural representation will hold implications for neural semantics, for the question of how neural representations can come to have content,

meaning, be about states of affairs beyond themselves (Maria, 2011).

## 6. Visuomotor Transformation

A specific but absolutely central topic to do with neural representation is visuomotor transformation, that is to say, how we use visual information to guide motor control. Here is the leading theory that we have two complementary visual systems, vision-for-perception and vision-for-action (Milner, 2011). They base their conclusion on a double dissociation between two kinds of disorder found in brain-lesioned human patients: visual form agnosia and optic ataxia. Milner claim that this functional distinction mirrors the anatomical distinction between the ventral pathway (to the side and near the bottom of the brain) and the dorsal pathway (to the rear and near the top of the brain) in the visual system of primates. Probably no other claim in cognitive neuroscience has attracted as much attention as this one in the past ten or twelve years. Another important body of work in visuomotor control focuses on the idea that spatial perception and motor output are interdependent. There are two broad approaches. One posits mental representations mediating between perception and action. This approach is often called representationalism. The other approach, a kind of antirepresentationalism, opposes this idea, arguing that intelligent, visually guided behaviour can be explained without positing intermediaries with representational or semantic properties between sensory input and motor output (Leach, 2008).

## 7. Colour Vision

The final two issues on which we will focus in this quick survey of issues currently at the interface between philosophy and neuroscience are colour vision and consciousness. Any complete theory of neural representation would have to contain a theory of both.

The biggest issue to do with colour vision, as we said, is the issue of how to think about the relationship of colour experience to the causal factors that produce colour experience. For example, experiences of different colours are the result of combinations of intensities of light of the three broad wavelengths of light to which the retina is sensitive (four wavelengths in some women) plus other factors. Light of three intensities at three wavelengths is nothing like redness as one experiences it. So how should we think of the relationship between the two?

Even worse, some argue that colour experience arises from processing that distorts the stimulus features that are its main causes, thereby largely constructing a world of perceived colour that has almost nothing to do with how the world actually is. For these people, perceived colour similarity is a systematic misrepresentation of the corresponding stimuli. How such systematic misrepresentation could have come to have a survival or reproductive advantage is just one of the puzzling, even baffling questions to which contemporary work in neuroscience on colour gives rise.

Most remarkably of all, we can have colour experiences that represent physically impossible colours. In a stunning example of neurophilosophy at work, Paul Churchland has shown that by exploiting shifts in experienced colour due to tiredness and habituation, experiences of colour can be brought about where the colours could not exist on standard colour wheels and other theories of the structure of colour and, moreover, would require physically-impossible states, for example, that things in one's world be emitting light and be emitting no light at the same time. Indeed, as Churchland shows, some of the colour experiences that we can have cannot even be represented by a colour sample (Churchland, 2008).

## 8. Consciousness

Most of the philosophical interest in consciousness starts from the question of whether consciousness could possibly be a physical process of any kind, let alone a brain process. A common view in philosophy of neuroscience is that everything to do with the mind, including consciousness, will turn out to be explicable in terms of neurophysiology—not even explanatory autonomy is allowed. If consciousness is not something that neuroscience can capture, then that hallowed shibboleth of neuroscience will be false and there will be at least severe limitations on the extent to which there could ever be a science of consciousness.

In the face of claims, at least something about consciousness is not neural or even physical at all, cognitive neuroscientists and their philosopher fellow-travellers have tended to one or the other of three different kinds of reaction:

- 1) They try to show that the claim is wrong (or incoherent, or in some other way epistemically troubled) (Dennett, 2005).
- 2) They just ignore the claim. This is the approach taken by many cognitive and neuro-scientists.
- 3) They throw science at it and attempt implicitly or explicitly to produce the kind of account that is supposed to be impossible

The usual way to argue the main idea in that there is nothing unique or *sui generis* about consciousness, is to tackle the arguments that claim that there is and try to show that they do not work. Here is a sample of such arguments. Dennett argued that because conscious experience is subjective, *i.e.*, directly accessible by only the person who has it, we are barred from ever understanding it fully, including whether and if so how it could be physical. For example, even if we knew all there is to know about bat brains, we would not know what it is like to be a bat because bat conscious experience would be so different from human conscious experience. Others extended this line of thought with zombie thought experiments and thought experiments about colour scientists who have never experienced colour (Dennett, 2005).

Law thought experiments are representative of the genre. Consider what philosophers call qualia: the introspectible aspects of conscious experiences, what it is like to be conscious of something. Those who hold that consciousness is something unique that there could be beings who are behaviourally, cognitively, and even physically exactly like us, yet they have no conscious experience at all. If so, conscious experience cannot be a matter of behaviour, cognition, or physical makeup (Law, 2010).

A variant, inverted spectrum thought experiments, urge that others could have radically different conscious experience of, in this case, colour with no change in behaviour, cognition, or physical makeup. For example, they might see green where we see red (inverted spectrum) but, because of their training, etc., they use colour words, react to coloured objects, and even process information about colour exactly as we do. If inverted spectra are possible, then the same conclusion follows as from the alleged possibility of zombies: consciousness is not safe for neuroscience. Weiskrantz and inverted spectra arguments strive to show that representations can have functionality as representations without consciousness. A more scientific way to argue for a similar conclusion involves appeal to cases of blind sight and inattentional blindness. Due to damage to the visual cortex, blindsight patients have a scotoma, a 'blind spot', in part of their visual field. Ask them what they are seeing there and they will say, "Nothing". However, if you ask them instead to guess what is there, they guess with far better than chance accuracy. If you ask them to reach out to touch whatever might be there, they reach out with their hands turned in the right way and fingers and thumb at the right distance apart to grasp anything that happens to be there (Weiskrantz, 2005).

Inattentional blindness and related phenomena come in many different forms. In one form, a subject fixates (concentrates) on a point and is asked to note some feature of an object introduced on or within a few degrees of fixation. After a few trials, a second object is introduced, in the same region but usually not in exactly the same place. Subjects are not told that a second object will appear.

## 9. Conclusion

In general, at the interface between neuroscience and philosophy at the moment, there is a great ferment. Results in neuroscience are shedding light on, even reshaping, traditional philosophical hunches about and approaches to the mind. And neuroscience is throwing up some new issues of conceptual clarification and examination of possibilities that philosophers are better equipped to handle than anyone else. We live in interesting times!

## References

- Akins, K. (2011). Of Sensory Systems and the "Aboutness" of Mental States. *The Journal of Philosophy*, 93, 337-372.  
<http://dx.doi.org/10.2307/2941125>
- Aldrich, C. (2006). Integrating Neuroscience, Psychology, and Evolutionary Biology through a Teleological Conception of Function. *Minds and Machines*, 6, 481-505.
- Almekhlafi, A. (2004). *Discovering Complexity: Decomposition and Localization as Scientific Research Strategies*. NJ: Princeton.
- Barta, J. (2002). *Philosophy of Science*. London: The Encyclopaedia of Cognitive Science.
- Bickle, J. (2012). Psychoneural Reduction: The New Wave. *Journal of Philosophy*, 78, 67-90.
- Birch, A. (2008). A Neurocomputational Perspective: The Nature of Mind and the Structure of Science. *Journal of Philoso-*



- phy, 76, 23-29.
- Bonk, C. J. (2000). Knowledge and Mind. *Journal of Philosophy*, 78, 30-44.
- Bowman, R. (2008). Eliminative Materialism and the Propositional Attitudes. *Journal of Philosophy*, 78, 67-90.
- Cassell, J. (2006). Towards a General Theory of Reduction. Part I: Historical and Scientific Setting. Part II: Identity in Reduction. Part III: Cross-Categorical Reduction. *Dialogue*, 20, 38-59.
- Churchland, P. M. (2008). *The Engine of Reason, the Seat of the Soul*. Cambridge MA: Rosenthal.
- Deane, R. (2001). Explanatory Pluralism and the Co-Evolution of Theories of Science. In W. Bechtel, & P. Mandik (Eds.), *The Churchlands and Their Critics* (pp. 17-47). Blackwell Publishers.
- Dennett, D. C. (2005). Why You Can't Make a Computer That Feels Pain. *Synthese*, 38, 415-456.  
<http://dx.doi.org/10.1007/BF00486638>
- Dretske, F. (2010). Knowledge and the Flow of Information. *Midwest Studies in Philosophy*, 430-444.
- Elwood, J. (2004). *Connectionism and the Mind: Parallel Processing, Dynamics, and Evolution in Networks*. Oxford.
- Fodor, J. A. (2010). Special Sciences (or: The Disunity of Science as a Working Hypothesis). *Synthese*, 28, 97-115.  
<http://dx.doi.org/10.1007/BF00485230>
- Funk, J. B. (2011). A Neurocomputational Perspective: The Nature of Mind and the Structure of Science. *Midwest Studies in Philosophy*, 512-534.
- Goos, M. (2007). *The Book of GENESIS*. New York: Nature of Mind, Oxford, 235-242
- Granič, A. (2007). Advertisement for a Semantics for Psychology. *Midwest Studies in Philosophy*, 10, 615-678.
- Hayes, B. (2009). What the Frog's Eye Tells the Frog's Brain. *Proceedings of the IRF*, 47, 1940-1951.
- Jumani, N. (2010). Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. *Journal of Physiology (London)*, 195, 215-243.
- Kan, S. (2011). Cooperative Learning Environment with the Web 2.0 Tool E-Portfolios. *European Journal of Social Science*, 21, 17-27.
- Kazu, I. (2011). An Investigation of Factors Affecting the Use of Educational Technology in Turkish Primary Schools. *Education*, 131, 510-524.
- Law, N. (2010). Policy Impacts on Pedagogical Practice and ICT Use: An Exploration of the Results from SITES 2006. *Journal of Computer Assisted Learning*, 26, 465-477. <http://dx.doi.org/10.1111/j.1365-2729.2010.00378.x>
- Leach, J. (2008). Do New Information and Communications Technologies Have a Role to Play in the Achievement of Education for All? *British Educational Research Journal*, 34, 783-805. <http://dx.doi.org/10.1080/01411920802041392>
- Lehky, S., & Sejnowski, T. J. (2011). Network Model of Shape-from-Shading: Neural Function Arises from Both Receptive and Projective Fields. *Nature*, 333, 452-454. <http://dx.doi.org/10.1038/333452a0>
- Luck, L. (2010). Maximizing the Usage of Technology-Enhanced Teaching and Learning of Science and Mathematics in English Program in the Malaysian Secondary Schools System. Online Submission.
- Maria, K. (2011). Teaching Art Using Technology: The Views of High School Students in Greece. *Review of European Studies*, 13, 98-109.
- Milner, A., & Goodale, M. (2011). *The Visual Brain in Action*. Oxford: Oxford University Press.
- Van Gelder, T. (2009). Mind as Motion: Explorations in the Dynamics of Cognition. *Journal of Consciousness Studies*, 5, 381-383.
- Weiskrantz, L. (2005). *Blindsight: A Case Study and Implication*. Oxford: Oxford University Press.