Journal of Cognition and Neuroethics

Pre-Conscious Noise

Bradley Seebach and Eric Kraemer
University of Wisconsin-La Crosse

Biographies

Brad Seebach is an Associate Professor of Biology at the University of Wisconsin-La Crosse. Seebach received his A.B. in English Literature from Cornell College and his Ph.D. in Neural Science from Brown University. He is a self-described developmental neuroscientist. Significant influences on his thinking come from years spent in the study of chemical engineering, linguistics, and computational neuroscience, in addition to many years of research and teaching in neurophysiology. His scholarship focuses on the development of neural systems. For many years, he has more specifically focused on the development of central pattern generator circuitry related to mammalian locomotion—one place where development may require unsupervised learning within a network of neurons, and yet be relatively accessible to analysis using the tools of an electrophysiologist.

Eric Kraemer is Professor of Philosophy at the University of Wisconsin-La Crosse. Kraemer received his A.B. in Philosophy from Yale University and his Ph.D. in Philosophy from Brown University. The overall goal of his current research program is to discover and defend a comprehensive naturalistic perspective by showing how the naturalist can rely not only on support from the advancing scientific program but can also successfully appeal to strategies developed and utilized by thinkers from other philosophical viewpoints, including the super-naturalist perspective. A convinced mind-body physicalist, Kraemer now views the newly emerging Hidden Properties Physicalism approach as providing the most helpful perspective for solving the mind-body problem, by explicitly supporting making connections between developments in the philosophy of mind and possibilities that current neurophysiology is making known.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). April, 2016. Volume 4, Issue 1.

Citation

Seebach, Bradley, and Eric Kraemer. 2016. “Pre-Conscious Noise.” *Journal of Cognition and Neuroethics* 4 (1): 89–112.

Pre-Conscious Noise

Bradley Seebach and Eric Kraemer

Abstract

Philosophers and neuroscientists take such different approaches to questions surrounding the existence or definition of conscious states that they often dismiss each others' viewpoints as irrelevant or meaningless. Yet there is a historically rich, informative interaction between philosophers and neuroscientists that we believe should be kept within the purview of each field. The traditions of philosophers who have considered the problem of consciousness, and the traditions of neuroscientists who have considered the problem of consciousness are examined in this discussion, in order to see how fundamental differences in the approach to the problem may produce a lack of obvious, common ground for discussion.

Reductive materialism and the dual-properties theory are presented as representative approaches employed by contemporary philosophers of mind, as they consider the problem of consciousness. On the side of the neuroscientists, we examine approaches to the same problem from perspectives used in neurophysiology and in computational neuroscience.

A key understanding emerges in the discussion, which is that introspective analysis of the mind cannot account for the emergence of apparently novel thought, and that this corresponds to the neurophysiologists' bane, which is the presence of apparently spontaneous neuronal activity, uncorrelated with identifiable patterns of behavior within a neuronal circuit, system, or whole animal. This "noisy", apparently spontaneous activity becomes of greater interest when viewed as a common weakness for both the philosophers of mind, for whom it represents the nothingness from which conscious states emerge, and for the neurophysiologists, for whom no sufficient tools exist to find correlated neuronal activity in a great many neurons or circuits in the majority of studies. The goal of discussion becomes to identify whether there is an unidentified precursor of a conscious state buried in the noise ("pre-conscious noise"), and then to identify a plausible explanation for what such a precursor might (roughly) look like.

In order to understand whether there are meaningful patterns within the noise, it is necessary to go beyond tools that are currently available to physiologists and consider what other demands are placed on the brain during behavior. Few modern neuroscientists would claim that a conscious thought that is in progress drives all activity in the brain—there are likely many, many patterns of activity in progress in the brain, some of which may be vying for attention. What rises to consciousness in the next moment of time may lie submerged in a partial pattern of activity that is not yet fully established, but that may become more fully established as the result of a blend of anatomical structure resulting from genetic and experience-driven development and ongoing learning, immediate sensory input, and existing or recent activity patterns in the circuitry of the brain. These are described as proximate and ultimate causes for behavior.

The oscillatory activity of the brain that appears as a fundamental drive for activity patterns and which is visible throughout life in electroencephalogram (EEG) measurements, varying in frequency range in a daily pattern, is considered as a driving force for the rise of partial patterns of neuronal activity to strong, completed patterns that could be correlated with behavior at either a conscious or non-conscious level.

The last piece of the discussion involves the necessity of providing an example of a pattern of activity that might have the qualities of a partial pattern that becomes a strong, completed pattern. For this, the distributed, content-addressable memory of a Hopfield network model is used. A neuronal network that uses a Hebbian-learning algorithm in an unsupervised manner is briefly examined, as it represents early learning for linguistically-relevant characteristics in a small network of mathematically simple, artificial neurons.

The resulting explanation, arising from principles of neurophysiology and computational neuroscience, is linked back into the philosophy of mind scholarship. "Hidden Natures Physicalism" offers a very interesting framework of consideration, within which our explanation of "conscious noise" or (perhaps more accurately) "pre-conscious noise" is found to have merit for explaining the essential characteristics of the mind-body problem that is at the heart of consciousness studies.

Keywords
Hidden Nature Physicalism, Computational Neuroscience, Content Addressable Memory, Resonant Patterns

Introduction

What is the problem of consciousness, and why is it so hard, both to get researchers outside of philosophy of mind to appreciate that there is a serious problem of consciousness, and, once clarified, to propose a program for how a solution might emerge? The solution proposed in this discussion is one that relates the disciplines of the authors of this paper, neurophysiology and philosophy. Philosophers, when they are thinking about the mind and mental phenomena, typically engage in first-person reflection, concentrate on the robust and varied qualities of the experiences they are having and then try to explain their natures and remarkable sensory and intentional features. Pains, itches and tingles, sensations of red, hallucinations and after-images, beliefs about non-existent beings, and desires for non-obtainable states-of-affairs all constitute typical subject matter for philosophers of mind contemplating consciousness. Neurophysiologists (and psychologists), on the other hand, engage in third-person observation, and, once they are confident with having established specific relations between types of mental states and events in the brain (or in the social context), then turn to postulating detailed physical (or social) mechanisms to explain how these states arise. For most of the past two decades, philosophers of mind have referred to the biologist's project as the "Easy Problem", while insisting that their own project is the "Hard Problem" (Chalmers 1995). How can the mechanisms of the neurophysiologist, wonderfully detailed though they are, account for the qualitative and intentional features of conscious experience? It is this lack of how a connection might be made between the two concerns that is often referred to as *the* "Explanatory Gap": exactly how is the itchiness of one's currently itching mosquito bite to be explained by the current structure and functioning of one's nervous system (Levine 1983)? But there is more to the story.

Reduction or Dualism?

Reductive materialism and the dual-properties theory are the two basic alternative approaches employed by contemporary philosophers of mind to account for consciousness. According to reductive materialists, all of reality is, at base, physical in nature, including conscious mental states, and all mental features, including consciousness, must somehow be reduced to, that is, completely explain in terms of, more basic physical features (Smart 1959, Armstrong 1968, Lewis 1966, Churchland, 1998, Hill 2009). This approach faces the problem just mentioned of elucidating how the basic entities and forces of physics and chemistry could give rise to the phenomenal or "felt" aspect of consciousness. This problem, therefore, seems better referred to as the Problem of the *Phenomenal* Explanatory Gap. This problem is neatly avoided by the dual properties theorist who postulates a special aspect of reality expressly to account for this felt gap between consciousness and the physical world.

Defenders of the Dual Properties Theory (DPT) come in three different varieties, the two most common versions of which are substance dualism and property dualism. The most important defender of the original version of substance dualism in the modern philosophical period was, of course, René Descartes (1648). Descartes claimed that the proper account of consciousness required postulating a second non-physical substance, a soul, in addition to the physical body. This version of the DPT has largely fallen out of fashion, being perceived as metaphysically non-parsimonious, and most current defenders of the DPT accept the simpler view that the features of consciousness, while non-physical, are nonetheless, features of a physical body. So, instead of there being two metaphysically different kinds of substances there is only one kind of substance with two metaphysically different kinds of properties. (See Nagel 1974, Campbell 1984, Jackson 1986, Levine, 1983, Chalmers 1995, Kim 2007.)

For the dual properties theorist, in order to account for the remarkable features of consciousness what is required is that we postulate, in addition to the basic materials of the physical sciences, a second, different aspect of reality to account for consciousness. The problem, of course, for defenders of the dual properties theory (DPT), is that they are equally challenged: they cannot explain [1] how this second aspect of reality arose, [2] how it appears to be caused by and causally influences physical reality, and [3] how this special aspect of reality actually creates consciousness. So, to be fair, defenders of the DPT need to admit that on their own view there remains an equally huge, if not larger "explanatory gap" between the two kinds of reality. Let us refer to this as the Problem of the *Metaphysical* Explanatory Gap: how can the physical world contain two different kinds of features, physical and non-physical?

There is also a third variant on the dual properties view, namely that of the non-reductive materialist. Proponents of this approach claim that while all of reality is ultimately physical, nevertheless there are some features of living beings, such as consciousness, which cannot be reduced to physicochemical properties (Cornman 1983, Searle 2004). Defenders of this approach face their own serious worries, namely [1] explaining how normal development from purely physical systems can give rise to two very different kinds of properties, (2) explaining how these two different kinds of properties are causally related to each other, and [3] explaining how physical features at one level, say the biological level, can themselves give rise to conscious experience at the psychological level. So again, there is an explanatory gap to be filled. Let us call this the *Level* Explanatory Gap: how can different, irreducible levels of physical reality arise? While it was once expected that functionalism combined with supervenience would provide the means to fill this gap, the current philosophical consensus seems is arguably that this approach has not delivered on its initial promise. (See Putnam 1999, Kim 2007; but compare Lewis 2004.)

From the perspective of scientific researchers who think that there may well be something of empirical value to be discovered from the scientific investigation of consciousness, however, all versions of the DPT are an unmitigated disaster. This is because defenders of all forms of the DPT place an impregnable metaphysical barrier between the philosophical and scientific explorations and explanations of the mind. Scientific researchers, on the other hand, continue to suggest that there must be further approaches that might be tried. And, some reductivist philosophers of mind would agree. In response we would like to suggest that the solutions of these two problems, the easy problem and the hard problem, might well go together. Why so? How is neuroscience research relevant to questions about the nature of consciousness?

The Relevance of Neuroscience

If we assume that humans are wholly biological beings, then at this point in the development of the history of science it seems that that neurophysiology, particularly when augmented with tools of computational neuroscience, may be well positioned to answer questions about how a human body can give rise to consciousness. Consciousness is a exceedingly difficult concept because it cannot be examined at a level that available scientific tools within any one discipline can manage. A synthesis of neurophysiology and computational neuroscience can, we believe, significantly advance our understanding. We will be investigating multiple causes for neuronal activity that appears to be correlated with behavior. We will also examine neuronal activity that appears to be uncorrelated with behavior and has therefore often been referred to as *noise*, or as being of random nature. As several classes of causal agents may exist for any neuronal activity, we will make use of a conceptual framework for the study of causation within the biological sciences established by Ernst Mayr (1961) and developed over the last half century by many others.

Neurophysiology provides tools that can examine nervous system activity that exists in the immediate context of a behavioral decision. These tools range from the extracellular observations of electrical activity in nerves that became common in the 19^{th} century, to 20^{th} century innovations such as the electroencephalogram (often credited to Hans Berger circa 1924) intracellular recording of action potentials (Hodgkin and Huxley 1939), and activity in ensembles of neurons making use of improvements in intracellular calcium-concentration imaging documented by Tsien (1980) in combination with advances in optics and in the handling of live nervous tissue specimens. Each of these techniques captures activity at a particular anatomical scale and time base of analysis, and is capable of finding correlated regularity in the midst of tremendous "noise" at the associated scale. Yet the analysis at the smaller anatomical scales is likely blind to emergent patterns of activity that appear in the analysis of larger circuits, and the larger-scale analysis is blind to regional sub-patterns that are likely to be of critical importance to attention selection and for (either) conscious perception or triggering reflexive responses. Each has its own effective time scale, also. Neuroscientists are, generally, quite aware of these limitations in scope.

Computational neuroscience can augment neurophysiology due to its ability to incorporate aspects of so-called "ultimate" causes derived evolutionarily. These include anatomical structure and mathematical rules that govern the interactions between neurons within a circuit. Such interactions will incorporate an individual's historical experience due to patterned sensory and contextual input through processes that can be broadly referred to as "learning", and will build upon the genetically-determined capabilities of the nervous system. In computational neuroscience, the experimental design of the sensory inputs and context can substitute for real-world experience that drives developmental change. These genetic and historical causes interact with immediate context to drive immediate decision-making processes in neuronal circuits that are responsible for recall and behavior. Consciousness sometimes has been attributed to the presence of synchronous, iterative activity in neuronal circuits and systems that produces strong learning in development phases and may trigger awareness thereafter (Llinàs et al. 1998). We will focus on the transition between (so-called) noise and stable patterns of activity that can be correlated easily with behavior.

The Conceivability Objection Considered

At this point some philosophers will object that all of these approaches to understanding consciousness are fundamentally flawed and for the same basic reason, viz. whatever discovery we claim to make regarding any such procedure that is then used to make the claim that the nature of consciousness is ultimately to be explained using that discovery we can then easily conceive of a creature who is just like the one upon which we are experimenting with the following difference: instead of being in a particular conscious state C^1 when the particular neurophysiological state or condition is present the creature either lacks consciousness altogether or is in a different conscious state, say C^2. From this thought experiment these theorists conclude that no neurophysiological account or explanation can fully explain consciousness (Nagel 1974, Campbell 1984, Jackson 1986, Chalmers 1995). But, we think that this argument, though often cited, is ill-considered. That there is yet no simple generally accepted neuroscience account yet of consciousness fails to prove that consciousness must be something other than neurophysiological activity. Consider this analogy. While physiologists before Harvey disagreed about how the blood circulated and could claim to refute their opponents' views with similar conceptual arguments, once Harvey established the correct pattern amidst the considerable physiological "noise" that persisted, all such contrary arguments fell by the wayside. We think that something similar is likely to happen with respect to consciousness, at some point.

A second useful analogy for the worry about consciousness is to be found in the debates between mechanists and vitalists shortly after the turn of the 20th century. The mechanists can be viewed as playing the role of the reductive materialists and the vitalists as playing the role of the dual properties theorists. The vitalists' (and later the organismalists') sole argument consisted in pointing to remarkable features in organic development which, it was claimed, just could not be explained by any purely mechanical process (Driesch 1908, Ritter 1919). Clearly, turn-of-the-20th-century mechanists had to be reductive mechanists, while their vitalist contemporaries were "dual property" vitalists. While the analogy may be comforting, in the meantime, how should one proceed?

What about "Noise"?

One promising answer counsels that we turn directly to "noise" itself. What is noise? We conceive of noise as a rich state of interaction in which activity patterns have not settled into resonance in the monitored anatomical region. Yet these suboptimal patterns of activity are continually influencing other, connected regions, offering momentary biases into patterned activity in those other regions. These momentary biases may combine in a manner that produces positive feedback, creating a resonant stable pattern, or they may fail to do so. So then, what does the noise represent? It represents multiple, potentially resonant patterns that are competing for attention or to trigger reflexive action. Also, for the neuroscientist, the noise is often what one tries desperately to filter out of one's results through either a hardware or software frequency filter or through post hoc statistical analysis, in order to reach a core pattern of some significance. One cannot easily publish uncorrelated data that cannot produce answers to useful questions!

Yet when we examine the nature of consciousness, the supposed unity of conscious experience, and the flow of attention from one focus to another, the noise may be precisely where we should be looking. In the period of time in which no recognizable, resonant pattern has yet been established, there is interaction occurring that involves evolutionary, developmental, and immediate causes for a behavioral decision. This interaction may produce strong influences that arise from each of these sources in a manner that masks a deterministic process. We believe it is therefore useful to consider causation more fully. And we need to use computational neuroscience in order to examine what might create "noise" while *en route* to stable, resonant patterns of activity.

Biological Causation, Including Evolution and Development

Mayr (1961) expands the Aristotelian view of causation into an explicitly "biological" framework, in which he describes proximate cause as things that a functional biologist might study. A neurophysiologist is our primary example of a functional biologist. Ultimate cause includes a broad variety of historical causes for a biological event, and includes the things that an evolutionary biologist might study. Mayr used an example of the proximate versus ultimate causes of a biological event such as the migration of a bird to illustrate his point. One of Mayr's great contributions was to provide an explicit distinction between these types of causal agents, and this has been an important contribution to the advancement of science (Beatty, 1994).

The apparent, attractive dichotomy between proximate and ultimate causes in biological science has been extended (perhaps appropriately) in a manner that ignores cautions that Mayr expressed (Laland et al. 2011), particularly with regard to the "muddle" of development (Amundson 2005). This "overextension" makes it more difficult to place developmental considerations into the proper framework with proximate and ultimate causes. In the words of Laland et al. (2011), "Mayr's proximate/ultimate distinction has proven problematic because it builds on an incorrect view of development that fails to address the origin of characters and ignores the fact that proximate mechanisms contribute to the dynamics of selection." Causation includes feedback loops that cannot be adequately described by a strict separation between proximate and ultimate causes.

We find it necessary, in the consideration of the neuroscience of consciousness, to carefully consider proximate causes for a behavioral choice, ultimate causes for that behavioral choice, and how developmental constraints and influences shape each of these. This is, essentially, a fundamental viewpoint in evolutionary-developmental approaches to biological science (Amundson, 2005). Because development incorporates the lifetime of experience of an individual into the background context of immediate decisions, we choose to include development as an ultimate cause for behavior, recognizing that distinctive time boundaries between ultimate and proximate causes are, in some cases, lost due to learning processes being capable of producing rapid change in neuronal circuitry under some circumstances. We find this to be an acceptable confusion, as it would only be applicable in earlier stages of development for many regions of the brain, as primary sensory regions and many of the non-cortical regions of the brain lose their plasticity after a critical period of development passes.

Neurophysiology tools can describe several layers of proximate cause (Calvin 1998, 2004), depending on the level of analysis and the tool being used. Computational neuroscience tools make explicit the constraints of memory storage and retrieval, and are therefore a required part of the consideration for how development (and incorporation of memory as a shaping influence on behavior) shapes both proximal and ultimate cause.

Conscious decision-making in a traditional sense may be thought of as a cause of behavior that would override ultimate causes, allowing a human to choose to ignore those ultimate causes that might arise through evolutionary shaping of the human body and nervous system, and to also ignore, at will, the developmental shaping of the body and its nervous system that is accomplished in preceding years through integral processes of learning. A strong bias towards this view likely develops from the natural border of levels of analysis – consciousness is limited by attentional focus and cannot be divided easily into multiple streams that can keep track of the minute contributions made by individual cells or even systems of cells (an anatomical and functional size barrier), nor can it be diverted from the behavioral time scale of seconds or tenths of seconds to analyze events occurring in only a few milliseconds or extending to minutes, hours, and days—these abilities are in the province of recorded language, which can bring context back to us in compressed or expanded time scales through the utilization of memory recall.

Consciousness and Oversimplification

Another strong bias towards the explanation that consciousness directs our activities comes from the evident creativity of conscious thought. However, consciously-directed creativity may be an incorrect explanation if there is lack of ability to understand underlying processes. Nate Silver (2012) describes a pattern for why humans consistently fail in necessary understanding for the accurate analysis of scientific, experimental results: a failure to eliminate personal viewpoint or bias from experimental design or analysis. We are, of course, blind to those things that we cannot see. Silver states that "we forget—or we willfully ignore—that our models are simplifications of the world." Biology is full of hidden patterns that have predictive power. Witness the rise of genetic explanations for human behavioral patterns, individual differences, and illness (Bargmann and Gilliam, 2013) that followed the discovery and developmental understanding of DNA. To which causal agents are we blind?

Calvin (2004) explicitly uses Mayr's proximate versus ultimate cause descriptions to describe layers of neurophysiological analysis from chemical-molecular events that determine a neuron's electrical state, to intercellular activity determined synaptically between two neurons, to a network of cooperating neurons, to the system of neurons and glial cells that govern a behavior. He compares this layering to an examination of the fibers that are stitched together in a pattern to make cloth, which then may be trimmed and organized to create an item of clothing. An example of a neuronal 'system' with such layering might be the visual system centers of the eye and brain, in which many neurons, made from similar components but exhibiting different morphologies and a variety of arrangements into networks, may together provide a sensory stream that can be integrated with a historical record of some kind maintained in memory, in order to produce an awareness of the visible world.

The Time Basis for Neural Circuits and Behavior

Less well-explored, perhaps, are the time bases for conscious decision-making, and the equally complex structuring of time. This is of great interest to us, because it offers an opportunity to explore what happens between the molecular and neuronal time scale of activity (nanoseconds to milliseconds), the behavioral time scale of hundreds of milliseconds to tens of seconds, and how this relates to observable chemical and electrical patterns of activity within nervous circuits. This is the realm of behavioral "noise" that has been evident to several generations of neurophysiologists, and which is generally ignored by others who make use of what the field of neurophysiology has been able to show at the more understandable, behavioral time scale. On a time scale, the proximate cause of a behavior may be an established, correlative pattern of activity among a large ensemble of neurons that represents a predictive state, as established by neurophysiologists and neuroscientists using EEG, fMRI, or other tools and generally detectable with a minimum duration of several hundred milliseconds or more. Yet that pattern of activity arose through many cycles of activity on a smaller time scale.

Those very brief cycles of activity have been viewed as 'noise' by many, but are they noise? Or do they represent a rich state of interaction that would be predictive of the final behavioral outcome, if we were not simplifying our model of analysis on both the time scale and the anatomical scale? If they have no purpose, then why do they exist as a nearly universal subscale activity across all orders of complex organisms and in simpler pattern generating neuronal circuits, such as those that drive locomotor activity?

The general problem may be thought of in this way. A complex pattern of activity in a highly-interconnected network of neurons exists in the nervous system at all times during a human life. The failure of such patterns is a common definition of death. The basic patterns reflect an anatomical substrate that is strongly influenced by evolution of the human body, but is also strongly influenced by interaction between evolutionary mechanisms and a lifetime of contextually-dependent change in the interconnections, and indeed, in the number and type of cells within many anatomically-identifiable brain regions. These are identifiably some of the ultimate causes for a behavior. Immediate context (proximate cause) produces a barrage of fluctuating inputs that drive change in the complex pattern of activity in this interconnected network, though the pre-existing state of activity and the developmental and evolutionary shaping of the neuronal and glial cell networks make the influence of immediate context dissimilar from one person to the next (Finn et al. 2015). In fact, because of the ability of the nervous system to reshape its connections through learning, the immediate context, if it reoccurs, may produce a dissimilar effect on the same person the next time.

Content-Addressable Memory

Computational neuroscience has described how a system of common elements in an interconnected network can store and recall memory patterns in a way that overlaps, so that no single memory trace is held discretely, yet each can be recalled as a form of "content-addressable memory". We will use the primary example of a Hopfield network (Hopfield 1982).

To illustrate the idea of content-addressable memory and its possible use in a neuronal circuitry, we will follow two paths of explanation. The first is to explain a general method of approximation that uses iterative processes of calculation, common to some types of mathematics and engineering methods, and which we believe to be a good analogy for a function of the recurrent, looping neuronal circuits that produce iterative activity in the brain. Such circuits retransmit information between regions of the cortex and the thalamus (and also involving various basal nuclei). The second path of explanation is to provide an example of useful information that could be distributed across a network of interconnected neurons to create overlapping memory in synaptic connections. The specific example we will use is to represent syllabic speech. Such sensory information may be involved in both memory formation and recall processes through a loop between cortical and subcortical auditory centers of the brain that can sustain iterative, auditory and cognitive processes and produce stable activity patterns.

Iterative Methods and Recurrent Neuronal Circuits

Iterative methods of approximation have a long history in mathematics. Characterizations of different, iterative methods are credited variously to Newton, Gauss, and many other mathematicians, and scientists. They were extended into common use in nonlinear systems in the computational sciences through the work of David Young (e.g., Young 1950; Kincaid et al. 2010) and others. The general idea is to use an equation that describes known parameters (representing boundary values) of a problem, then insert into the equation a hypothesized, possible value x. Solving the equation using the hypothesized value returns a closer approximation to the real value of x. This new value can be substituted back into the same equation, and reiteration of the calculation will produce a new, even more accurate approximation to the real value of x. This gradual process of approximation is often referred to as a "relaxation" process when it is used to identify one or more stable states for a complex system, in which values may represented in two- or three-dimensional arrays, and for which the iterative mathematical techniques are often drawn from linear algebra.

We will use the term “relaxation” to describe the iterative approximation techniques that are used in computational models and that we believe are employed in the nervous system for purposes of memory recall and selection of behavioral responses. The term “relaxation” is derived from original conception of a reduction in chaos of a highly-variable system (in terms of energy) to a more coherent, stable energy state. In a Hopfield network, if part of a pattern that has been stored in content-addressable memory is presented as an input to the network (of artificial neurons), those neurons are excited in a manner that will return a closer approximation of the stored pattern. This closer approximation can be re-presented as an input to the network, and the following output should produce an even better representation of the stored pattern. Hence, when used in an iterative fashion, the recall process gradually improves the result as it “relaxes” to a coherent, pattern stored in memory, and the input and output of the network come to match one another more and more closely – the minimization of differences is thought of as representing a minimal, more “relaxed” energy condition as compared to the initial response state.

A system of 100 neurons with simple interconnections in a Hopfield network is estimated to be capable of storing without error up to 15 memory traces (Hopfield 1982) in which the state of each neuron is important to the correct learning and recall of each of the 15 memory traces. Storing additional memory traces produces “errors” in recall due to pattern interference that allows confusion of patterns. This type of error is actually of great interest to our discussion as well – for the confusion of one pattern with another may be deemed an error or it may be the basis of fruitful, creative processes. It is worth noting that the Hopfield network can produce an error in recall if the initial input either does not match any of the stored “memories” or if it is indeterminate between two or more stored memories.

Many connectionist models have been built upon the ideas represented in the Hopfield network combined with variations on the concept of Hebbian learning (Hebb 1949), and we briefly examine one such use.

Recurrent Neuronal Circuits and Oscillatory Activity

Iterative, synchronous activity in the mammalian brain has been identified in thalamic neurons that project to many regions of the cortex (Hunnicutt et al. 2014). Synchronous, oscillatory neural activity in the gamma-wave frequency range has been postulated to be critical to binding together neural centers that work together for behavioral purposes in development, consciousness and attentional modulation (Llinas et al. 1998, Miltner et al. 1999) using mechanisms of Hebbian learning (Hebb 1949, Caporale and Dan 2008) and accretion of neuronal activity that shapes alpha-wave oscillatory activity (Bollimunta et al. 2011). For our purposes, it is useful to note that gamma-wave frequency range is 40-100 Hz, offering the possibility of many iterations of activity, each producing greater synchrony in neuronal activity to perhaps cross a threshold to produce behavior, within the time frame of several hundred milliseconds that is associated with even rapid, primed voluntary behaviors in response time experiments. Oscillatory brain wave activity in the slower beta-wave frequency range (12-40 Hz) is associated with awake, active mental behavior that corresponds with conscious thought, and may represent a more standard pace of "iterative approximation" that human brains could use for complex processes—in which current sensory conditions and internal processes might be matched with memory in order to produce recall and to choose an appropriate action.

In studies of "readiness potentials" (Libet et al. 1982, 1983) that precede voluntary activity, a gradual increase in synchronicity of neuronal activity is noted and reaching a threshold level of excitatory synchronicity seems to be associated with awareness (Mathewson et al., 2009). A useful, stochastic "accumulator model" that links readiness potentials (represented in EEG traces) with behavioral tasks has been published by Schurger et al. (2012).

The type of information that the nervous system needs to learn to work with, and to match during a process of recall, can be represented with a matrices of numbers that represent excitatory activity in synaptic connections that represent an orderly, topographic map of a type of information. Our example will use auditory information and basic assumptions about peripheral auditory processing used to create a model of prenatal speech perceptual development (Seebach et al.1994). Speech samples that were sufficient for the development of perceptual discrimination of elementary consonant-vowel syllables, differing only in the initial stop consonant's place of articulation, have acoustic energy intensities represented in different frequency bands in the range of human hearing, spread across a brief period of time (Figure 1A). Sounds such as these, presented repeatedly as through the apparatus of the auditory system to an interconnected group of artificial neurons (see Bienenstock et al. 1982) whose connections (representing synapses) change in a Hebbian learning process, will shape responsive patterns of those neurons (Figure 1B) in a manner that produces discriminative ability (Seebach et al.1994). Different syllables can be discriminated by the presence or absence of acoustic energy and energy transitions at specific times and frequency regions. The purpose of the particular study was to show that this type of discrimination could be learned by a neuronal circuit in an "unsupervised" manner—simply as a matter of experience, with no "teacher" or confirmation of right or wrong responses—in other words, a learning process that could explain developmental changes that can occur very early, prior to the time of full, social interaction.

This is a very limited example of a memory process that could incorporate aspects of ultimate causes such as evolutionary shaping of the initial, anatomical network of neurons and the historical, developmental shaping of regions that become involved in more specific processing tasks. Iterative calculations representative of oscillatory brain wave patterns can be used to aid the development of the memory traces. For example, in the Seebach et al. study mentioned earlier, several thousand presentations of the CV syllable stimuli were needed in order to produce stable processes of learning. At first, this seems to be a rather high number of presentations—but the iterative, oscillatory circuits of the human brain would greatly reduce the number of real-world presentations that might be needed. The brain supplies many, additional iterations of each behavioral presentation. This amplification of learning and memory could be a reason for the natural, oscillatory processes noted by Llinas et al., (1998), Miltner et al., (1999), Bollimunta et al., (2011) and many others. Oscillatory patterns provide robust, intensifying patterns that can underlie Hebbian learning, and also could account for observed intensification of readiness potentials. It is perhaps unsurprising, then, that each human being appears to have a unique "neural fingerprint" (Finn et al., 2015), given that ultimate causes for brain activity and decision-making will be unique for each individual, given unique genetic heritage (excepting identical twins) and a unique life-time of experience for each. Within the 'noise', therefore, are unique sub-patterns of activity that may produce different responses across individuals or within an individual as context changes.

Conscious Noise and Hidden Nature Physicalism

What we find both interesting and promising about this approach to consciousness is that it indicates how one might get past the current impasse in philosophy of mind with regard to how the reductive materialist project should be pursued. It should be noted that, just as there are three versions of the DPT, there are at least two versions of reductive materialism (RM) as well. Some versions of RM make explicit claims regarding which physical features or mechanism with which we are to identify mental states, and some do not. Let us call the former Explicit Reductive Materialism or (ERM) and the latter Hidden Nature Physicalism. Early-twentieth century attempts by behaviorists to account for mental states solely in terms of behaviors and behavioral tendencies would count as one version of ERM. Mid-twentieth century thinkers then proposed to identify mental states and their features with specific neurophysiological states and their features (a view known as the "Identity Theory"). This is also a version of ERM. (Smart 1959) This view soon had to confront the conceptual objection that one could not expect exact neurophysiological identity between two individuals in the same mental state only similarity, and such similarity needed to be explained as functional identity. (Fodor 1981) There now seems to be empirical confirmation for this possibility. (Finn et al. 2015]

The latter part of the twentieth century was then dominated by attempts by functionalists to account for mental phenomena in terms of what sorts of functional roles particular physical states might instantiate.[1] But, while initially enormously promising, the Functionalist Turn has not proven able to deal successfully with the Hard problem. (Ludlow, et. al., 2004) The inability to produce a generally convincing functional account of consciousness has led some recent philosophers of mind to embrace instead a non-explicit version of reductive materialism which is now often referred to as "Hidden Natures Physicalism". But, what is Hidden Nature Physicalism and why does the program for the study and discovery of the physical nature of consciousness here proposed involving 'Noise" seem to support it?

Hidden Nature Physicalists claim that the nature of mental states is not completely revealed in our experiences of them, which entails that that there is more to consciousness than its experiential nature. For example, Chris Hill says the following:

> "If the Cartesian argument is to succeed…the essential nature of pain must be fully accessible to us when we experience pain. But it is precisely this thesis about the essential nature of pain that is called into question….There is no guarantee that experiential representation of pain will do full justice to its essential properties." (Hill 2009, 118)

Similarly, Patricia Churchland comments:

> What is troublesome is the idea that all the reality there is to a sensation is available through sheerly having it….I suggest, instead, a rather simple alternative: A sensation of pain is real, but not everything about

1. Whether Functionalism is itself a version of Reductive Materialism or a form of Nonreductive Materialism is not a settled matter. Some functionalists consider themselves to reductivists, other regard themselves as non-reductivists. The key to the disagreement is which and how strict an account of reduction one accepts. Those who accept Ernst Nagel's account of reduction will count themselves as non-reductivists; those inclined to the Kemeny-Oppenheim approach will consider themselves to be reductivists. (Nagel 1961, Kemeny and Oppenheim 1956)

> the nature of pain is revealed in introspection—its neural substrate, for example, is not so revealed." (Churchland, 1998, 117).

Similarly, Michael Tye says he rejects the view that "experience itself reveals red or canary yellow as simple, as not having a hidden nature". (Tye, 2009, 142). Thus for Hidden Nature Physicalists there must be something in addition to mere experience to account for consciousness, and this something is its physical nature, about which we are currently ignorant.

Why is HNP difficult for some physicalists to accept? Many physicalists are reluctant to make claims defending their views that are ultimately based upon current ignorance and the issuing of empirical promissory notes. Such a strategy seems not to be much an improvement over the various forms of the DPT discussed earlier. But, even a casual study of current brain research makes one painfully aware of the woeful state of our ignorance regarding how the brain might work to produce consciousness. If such ignorance really is our current epistemic condition, then, we should act accordingly. If we take the objections to all forms of DPT to be conclusive, and, if we think that objections to all other forms of RM are also convincing, and if we think that some version of physicalism is the only palatable alternative, then, as distasteful as some may find it, the most reasonable alternative seems now to be to embrace HNP.

While variants on HNP have been proposed in the past, these variants typically involved claiming either that there was a conclusive epistemological limitation that prevented humans from ever figuring out what physical properties consciousness might be,(McGinn, 1989) or that consciousness was to be conceived as analogous to anti-matter and could not be studied scientifically (Chalmers, 1995). Both of these alternatives, unlike the sort of approach embrace by the conscious noise proposal defended here, are seriously unsatisfying from the perspective of the scientific researcher.[2]

Let us take HNP, then, simply to be the general thesis that [1] although humans are wholly physical beings with no special levels, messily emergent features or spooky stuff, simple mind-brain identity theories are inadequate to account informatively for all mental phenomena; [2] functional accounts, while somewhat helpful in accounting for similarities across individuals and species, are nonetheless unable to account in an enlightening way for phenomena such as consciousness, and [3] there is some, as yet, undiscovered purely physical aspect of human existence that does explain such

2. McGinn's Mysterianism seems under-motivated and sells comparative scientific procedures short, (Kraemer 2007); and the Chalmers' anti-matter approach deliberately flirts with pan-psychism, a view which seems evolutionarily under-motivated as well as unsupported by our best current evidence.

phenomena. It should be clear from what has been said that there are at least two version of HNP, those variants which claim that, for one reason or another, mental states are physical but their physical natures are not discoverable by human beings, and those which urge that for all we know we may very well be able to discover and adequately explain the physical nature of consciousness. Let us call the first variant, Forever Hidden Nature Physicalism, (FHNP), and the second variant Currently Hidden Nature Physicalism, (CHNP). From the perspective of scientific researchers working on consciousness the latter variant is the more interesting, and the one which relates directly to the noise account proposed above.

What the pre-conscious noise account actually makes plausible is why the specific nature of consciousness should have remained hidden, and why philosophers of mind and other interested parties should be happy to support neuroscientists in their efforts to try to make clear just what sorts of developed patterns different forms of consciousness might turn out to be. And this is precisely what the pre-conscious noise proposal does: [1] it suggests a source for items out of which consciousness might be constituted; [2] it suggests a mechanism or physical procedure whereby consciousness might be seen to develop over time; and, [3] it also makes it very clear why the kind of enlightening account of phenomenal features that we think is currently missing from the physicalist camp has not been readily apparent to all; which [4] in turn explains explicitly the current and contingent hiddenness required by the CHNP. These are, we think, very serious advantages that the CHNP offers with respect to solving the problem of consciousness.

Conclusion

The conscious noise approach to consciousness outlined here promises a satisfying reply to a requirement set down some years ago by Donald Davidson with respect to the mind-body problem. (Davidson 1980) Davidson claimed that any adequate solution to the mind-body problem must be able to provide not only a convincing account of mental phenomena, but must also be able to provide a convincing explanation as to why it took so long for human beings to figure out how to answer the mind-body problem. And, the pre-conscious noise proposal is certainly able to provide a most satisfactory explanation on that score.

So, one might then ask, if the project of determining the mechanics of consciousness are to be turned over to the neurophysiologists and their scientific allies, what role should philosophers of mind who support the CHNP continue to play? Another way to pose the same question is to ask: what intellectual burdens must CHNNP defenders assume?

There seem to be at least three. First, defenders of CHNP need to provide convincing arguments for to articulate the numerous advantages of the CHNP approach. Second, there are serious objections that have been raised specifically against this view (Fumerton 2013, Robinson 2014), and defenders of the CHNP approach need to provide convincing philosophical rebuttals to them, as well as trying to anticipate and respond to other philosophical worries for CHNP that are likely to be raised. And, third, supporters of CHNP further need to be ever on the look-out for a plausible empirical strategy that supports the view by demonstrating not only how consciousness might arise but also by indicated how the currently hidden nature of consciousness might come to be made public. This discussion has attempted to provide some initial support for all three of these projects.

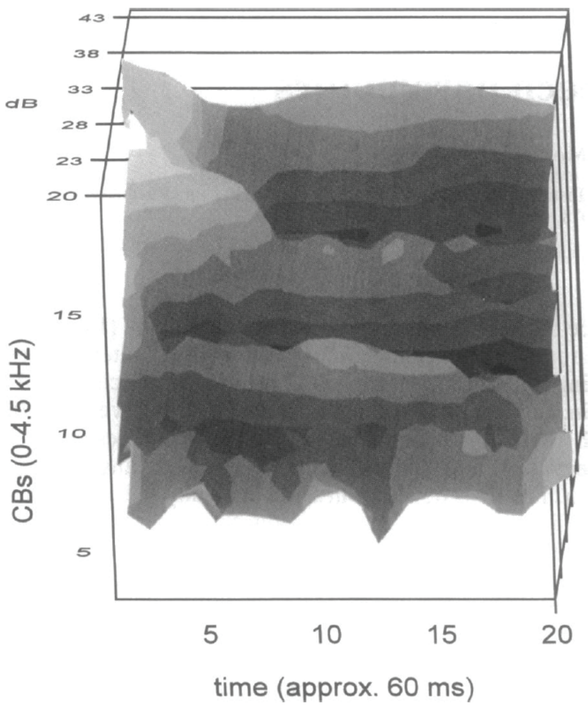

Figure 1. Modified from Seebach et al. 1994. (A) Average energy contour for each of the three training syllable types for speaker BR, shown as gray-scale images. The lighter areas of these images represent the presence of greater acoustic energy, with the ordinates of each image representing increasing frequency on a critical-band scale, and the abscissa of each image representing increasing time, which goes from 0 to 38 ms as marked by the start of each sampling window. (B) Gray-scale images of resulting synaptic weights for the five cells (neurons) of a BCM artificial neuronal network following developmental training using the inputs shown in (A). Collective responses provide a clear ability to discriminate among the different training syllables, and among the same syllables as spoken by other people.

Philosophy References

Armstrong, David. 1968. *A Materialist Theory of the Mind*. London: Routledge.

Campbell, Keith. 1984. *Body and Mind*, 2nd ed. Notre Dame, IN: Notre Dame University Press.

Chalmers, David. 1995. "Facing Up to the Problem of Consciousness." *Journal of Consciousness Studies* 2: 200–219.

Cornman, James. 1981. "A Non-Reductive Identity Thesis about Mind and Body." In *Reason and Responsibility*, edited by Joel Feinberg, 285–296. Belmont, CA: Wadsworth Publishing Co.

Churchland, P.S. 1998. "Brainshy: Nonneural Theories of Conscious Experience." In *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*, edited by Hameroff et al., 109–126. Cambridge, MA: MIT Press/Bradford.

Davidson, Donald. 1980. "The Material Mind." In *Essays on Actions and Events*, edited by Donald Davidson, 245–61. Oxford: Clarendon Press.

Descartes, Rene. 1648. *Meditations on First Philosophy*.

Driesch, Hans. 1908. *The Science and Philosophy of the Organism: Gifford Lectures delivered at Aberdeen University, 1907-1908*. London: Adam and Charles Black.

Fodor, Jerry. 1981. "The Mind-Body Problem." *Scientific American* 244: 114–225.

Fumerton, Richard. 2013. *Knowledge, Thought and the Case for Dualism*. New York: Cambridge University Press.

Hill, Christopher. 2014. *Meaning, Mind and Knowledge*. Oxford: Oxford University Press.

Jackson, Frank. 2004. "Post Scripts." In *There's Something about Mary*, edited by Peter Ludlow, Yugin Nagasawa and Daniel Stoljar, 407-444. Cambridge, MA: MIT Press.

Jackson, Frank. 1986."WhatMary Didn't Know." *Journal of Philosophy* 83: 291–295.

Kemeny, J. and Paul Oppenheim. 1956."On Reduction." *Philosophical Studies* 7: 6–19.

Kim, Jaegwon. 2007. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press.

Kraemer, Eric.2006. "Moral Mysterianism." *Southwest Philosophy Review* 22: 69–77.

Levine, Joseph.1983. "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64: 354–361.

Lewis, David. 1966. "An Argument for the Identity Theory." *Journal of Philosophy* 63: 17–25.

Ludlow, Peter, Yugin Nagasawa, and Daniel Stoljar (eds.). 2004. *There's Something about Mary.* Cambridge, MA: MIT Press.

McGinn, Colin. 1989. "Can We Solve the Mind-Body Problem?" *Mind* 98: 349–366.

Nagel, Ernst. 1961. *The Structure of Science*. New York: Harcourt, Brace and World.

Nagel, Thomas. 1974. "What is it like to be a Bat?" *The Philosophical Review* 83: 435–450.

Nagel, Thomas. 2012. *Mind and Cosmos*. Oxford: Oxford University Press.

Putnam, Hillary. 1999. *The Threefold Cord: Mind, Body and World*. New York: Columbia University Press.

Robinson, William. 2015. "Hidden Nature Physicalism." *Review of Philosophy and Psychology* 7 (1): 1–19.

Ritter, William. 1919. *The Unity of the Organism, or the Organismal Conception of Life*, two volumes. Boston: Richard G. Badger.

Searle, John. 2004. *Mind*. Oxford: Oxford University Press

Smart, J.J.C. 1959. "Sensations and Brain Processes." *The Philosophical Review* 68: 141–156.

Tye, Michael. 2009. *Consciousness Revisited*. Cambridge, MA: MIT Press.

Biology / Neuroscience References

Amundson, R. 2005. *The Changing Role of the Embryo in Evolutionary Thought: Roots of Evo-Devo* (Cambridge Studies in Philosophy and Biology). Cambridge: Cambridge University Press.

Bargmann, C.I., and T.C. Gilliam. 2013. "Genes and Behavior." In *Principles of Neural Science*, 5th Edition, edited by Kandel et al., 39–65. New York: McGraw Hill.

Beatty, J. 1994. "Ernst Mayr and the Proximate/Ultimate Distinction." *Biology and Philosophy* 9: 333–356.

Bienenstock, E.L., L.N Cooper, and P.W. Munro. 1982. "Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex." *Journal of Neuroscience* 2: 32–48.

Bollimunta, Anil, Jue Mo, Charles E. Schroeder, and Mingzhou Ding. 2011. "Neuronal mechanisms and attentional modulation of corticothalamic alpha oscillations." *Journal of Neuroscience* 31 (13): 4935–4943.

Calvin, W.H. 1998. "Competing for Consciousness: A Darwinian Mechanism at an Appropriate Level of Explanation." Journal of Consciousness Studies 5: 389–404.

Calvin, W.H. 2004. *A Brief History of the Mind: From Apes to Intellect and Beyond*. Oxford: Oxford University Press.

Caporale, Natalia, and Yang Dan. 2008. "Spike Timing–Dependent Plasticity: A Hebbian Learning Rule." *Annual Review of Neuroscience* 31: 25–46.

Finn, E. S., X. Shen, D. Scheinost, M. D. Rosenberg, J. Huang, M. M. Chun, and R.T. Constable. 2015. "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity." *Nature Neurosci*ence 18 (11): 1664–1671.

Hebb Donald. 1949. *The organization of behavior*. New York: Wiley & Sons.

Hodgkin, A.L., and A.F. Huxley. 1939. "Action potentials recorded from inside a nerve fiber." *Nature* 144: 710-711.

Hopfield, J.J. 1982. "Neural networks and physical systems with emergent collective computational abilities." *Proc. Natl. Acad. Sci. USA* 79: 2554–2558.

Hunnicutt, B.J, Long, B.R, Kusefoglu, D., Gertz, K.J., Zhong, H., and T. Mao. 2014. "A comprehensive thalamocortical projection map at the mesoscopic level." *Nature Neuroscience* 17: 1276–1288.

Kincaid, D.R., R.S. Varga, and C.H. Warlick. 2010. "The life and times of Dr David M. Young, Jr." *Numer. Linear Algebra Appl.* 17: 743–757.

Laland, K.N., K. Sterelny, J. Odling-Smee, W. Hoppitt, and T. Uller. 2011. "Cause and Effect in Biology Revisited: Is Mayr's Proximate-Ultimate Dichotomy Still Useful?" *Science* 334: 1512–1516.

Libet, B., C.A. Gleason, E.W. Wright, and D.K. Pearl. 1983. "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act." *Brain* 106: 623–642.

Libet, B., E.W. Wright Jr., and C.A. Gleason. 1982. "Readiness-potentials preceding unrestricted 'spontaneous' vs. pre-planned voluntary acts." *Electroencephalogr Clin Neurophysiol* 54: 322–335.

Llinàs, R., D. Contreras, and C. Pedroarena. 1998. "The neuronal basis for consciousness." *Phil.Trans. R. Soc. Lond. B* 353: 1841–1849.

Mathewson, K.E., G. Gratton, M. Fabiani, D.M. Beck, and T. Ro. 2009. "To See or Not to See: Prestimulus α Phase Predicts Visual Awareness." *Journal of Neuroscience* 29: 2725–2732.

Mayr, Ernst. 1961. "Cause and Effect in Biology." *Science* 134: 1501–1506.

Miltner, W., C. Braun, M. Arnold, H. Witte, and E. Taub. 1999. "Coherence of gamma-band EEG activity as a basis for associative learning." *Nature* 397: 434–436.

Schurger, A., J. Sitt, and S. Dehaene. 2012. "An accumulator model for spontaneous neural activity prior to self-initiated movement." *Proc. Nat. Acad. Sci. USA* 109: E2904–E2913.

Seebach, B., N. Intrator, P. Lieberman, and L.N Cooper. 1994. "A model of prenatal acquisition of speech parameters." *Proc. Natl. Acad. Sci. USA* 91: 7473–7476.

Silver, Nate. 2012. *The signal and the noise: why so many predictions fail—but some don't*. New York: Penguin Books.

Tsien, R.Y. 1980. "New calcium indicators and buffers with high selectivity against magnesium and protons: design, synthesis, and properties of prototype structures." *Biochemistry* 19: 2396–2404.

Young, D.M., Jr. 1950. *Iterative Methods for Solving Partial Difference Equations of Elliptic Type*. Ph.D. thesis. Cambridge, MA: Harvard University, Mathematics Department.