

Copyright Michael A. Fuller 2013

This material is in copyright. Subject to statutory exception,  
no reproduction of any part may take place without the  
written permission of Michael A. Fuller, (That should scare them!)

# The Neuronal Self

*The Embodiment of Experience and Selfhood  
in the Neural Matrix of the Mind*

Michael A. Fuller  
University of California, Irvine  
August 14, 2001

In southern California, where I live, the snails come out to forage at dusk and retreat into the shade at dawn. As I watch their slow, purposeful march toward my vegetables, I wonder how they understand the world, how they put it together. Is their “world” three-dimensional; is there an “up” as well as forward and sideways? After all, they spend their lives navigating surfaces and need simply to account for the ease or difficulty of their travel along their routes. The idea of an extra dimension would help—as it does in the modern account of gravity—but it’s not really needed. Still, I wonder how the snails might explain their experience of flight when I pluck them up and toss them over the fence into the ivy. Sheer terror? A moment of ecstasy? An overthrow of all snail philosophy?

Snails presumably have no philosophy. Yet imagining such an object as a metaphysical snail has its uses: even if we were to grant that reflective musings lie behind the waving horns of the snail, we surely would reject those thoughts as paltry things. Nonetheless, the point of this small book is to suggest that as creatures knowing the world through the constraints of the body, there really isn’t that much difference between us and the snails in my garden. More importantly, however, being a biological entity—living through a body and a brain—does not diminish the astonishing richness and complexity of our encounter with the world.

## Chapter One

### A Biological Imagination

#### Generalities

Biology, like geology, is an empirical science. Geology explores the patterns of transformation in rock. Biology explores those patterns for living organisms. In both cases, the patterns are experiential and phenomenal rather than ontological or metaphysical. A column of basalt is geological; a human is biological. In this sense, being biological is a way of approaching the world of experience: ingesting objects, extracting energy, shrinking from harm or overcoming it, and so on. Being biological is not really a problem: we are born to it. The difficulty is in understanding what being biological might mean and what it does *not* mean.

As biological entities, our empirical encounter with the world has a microstructure that goes deep. Neuroscientific inquiry increasingly reveals to us how synaptic structures mediate human experience. Yet while systems subserving perception, cognition, emotion, and judgment have much to say about the experience of listening to Beethoven or reading Du Fu, these biological structures help to articulate our qualitative responses rather than to replace them. Understanding this issue—how knowledge of the biological microstructures of perception, cognition and emotion help articulate the meaning of lived experience—is a difficult and delicate project. It contravenes many habits of

thought and in the end requires an imaginative leap to reconceive our manner of engagement with the world.

### **Turning Inward**

We readily grant that a snail is biological, for there is not much to the creature. We even readily grant that our bodies are biological. Consider any part of that body: the fingers that plucked up the snail, for example. Fingers suffer myriad indignities that remind us of their physicality. They become raw, red, and stiff with the cold; they get splinters, thorns, and hangnails. When cut, they bleed. And yet, to make the inevitable turn in perspective, the fingers know nothing of these matters: we do. We know our fingers' color and shape: retinal information makes its way through the thalamus and superior colliculus, and somewhere in the higher visual cortex the data streams reunite shape information with color, and a little higher up in the temporal lobe, these in turn activate what we know as "fingers."

These "fingers" are not a simple coding but a complex web. The attentional system gives our digits high priority because very early in childhood (though never early enough, from a parent's point of view) we learn to watch out for them. Sucking them is soothing, but getting them "dirty" makes them taste funny and brings in grit and fuzz. They get banged, stepped on, and pinched. We put them in places they should not go. That is, they have a complex history captured in the structure of their synaptic representation: they are not a sheer unmediated physical presence.

We cannot approach the body and body parts without interpretation. We take sensation as immediate and physical, yet where in our sensations does the body end and the mind begin? I return to that errant snail with its grey-brown horns waving at me. Synapses from the retinas of my eyes send spikes of information to the thalamus and superior colliculus and then on to the visual cortex, but where does the quality of “slimy grey-brown” emerge? When does the pattern of spikings become the mental entity “snail?” For that matter, is there even any point to referring to distinctly mental—as opposed to biological—entities? No matter where we look for unmediated physicality, we confront an inward spiral from body to mind. The body is mediated by the concept of the body. Sensation through some inward alchemy shifts from biological innervations to mental experience. And “we” who experience the shocks to the body are lodged deep within the “Cartesian theater” of the mind.<sup>1</sup> To force ourselves out of that theater, we need to pursue this inward trajectory, to see the ways in which biological structures and processes inform the mind and the self who watches, experiences and responds.

What is at stake here is more than just classical questions of biology. Pushing biological processes inward—into the realms of language, thought, and selfhood—addresses problems of much broader importance about the empirical *meaning* of experience. By articulating the pathways and structural logic of such functions as perception and semantic memory, neuroscience begins to explain not just the processes that subserve experience but the content as well. As I look at the

---

<sup>1</sup> The phrase is Daniel Dennett’s from *Consciousness Explained* (Boston: Little, Brown and

snail held between my fingers, I do not actually feel the snail: my brain just does the best it can to report the information from sensors on the fingertips. Nor do I see the snail: no little me looks directly out from the windows of the eyes. Instead my visual awareness of the snail is a conversation between memory structures—everything from simple shape recognition to phonological recall (the sound “snail”) to the web of associations about vegetables, snail bait, and the like—with the retinal information as transformed through the filtering system of the visual cortex. What I think about the snail I think with and through these structures and processes. Yet this thinking through biological forms is true not only of reflecting on the lowly snail but also of watching my daughter play with our cats, of my own profession of studying and teaching Chinese poetry, and indeed of all empirical experience.

Delineating essentially continuous causal processes that communicate between the outside—the body and the environment—and the internal states (i.e. the brain structures and activities that embody what we think, feel, learn, and recall) does away with much of what we think of as the mind-body distinction. Much of what had been the *res cogitans* gets swallowed up by the *res extensa*. Precisely because neuroscience offers a powerful empirical theory of reference (i.e., a theory of the relationship between internal signifiers and external objects, which includes a theory of language), we must be cautious. Much is at stake, not just for neuroscience but for the humanities and all other disciplines of human self-understanding: it is important to frame this empirical theory of reference properly.

### **Grounding Knowledge of Objects: Mind, God, Reason, and Genius**

That biology imposes constraints on knowledge is not a new idea. Both Plato and Aristotle confronted the limits of the senses and shaped arguments about the sort of knowledge we can derive from the body and the sort that we cannot. In order to defend the possibility of sure, stable knowledge, Plato concluded that the world accessible to the senses did not matter so much as the realm open to the mind (nous). The forms and structures of sensory experience were but uncertain shadows. European culture has been debating this issue ever since. Christian theologians, Renaissance humanists, Enlightenment philosophers and Romantic poets all approached Plato's dualistic framing of the question of knowledge, experience, and the body with their own commitments. Each development added layers to our cumulative understanding of how the senses open onto a world of experience and how the mind determines meaning; and no layer ever fully replaced the earlier strata. In our present society, some live with Christian commitments, others with the humanist's fascination with the world or with the Romantic's belief in the imagination. The neuroscientific model of the mind is an addition to the debate. Because it explains the empirical structuring and processing of the content of the human engagement with the phenomenal world, it cannot be entirely neutral in its account. It potentially clashes with other models to the extent that they offer causal interpretations of the meaning of phenomenal experience. A brief survey of the traditions that contribute to our current understanding of how the mind reads



the world will help clarify the nature of the conflicts and the stakes involved in a biological approach to experience.

The central problem of experience and knowledge famously embodied in Plato's image of the cave is that the world of senses seems unstable and insubstantial. The possible objects of knowledge arise, transform, decay, and vanish. At the beginning of the Western philosophical tradition, Plato proposes that what we see in the phenomenal realm are no more than fleeting shadows of unchanging entities: the Platonic Ideas. The problem is that we ourselves are caught in the world of transformation, so how do we have access to these metaphysical Ideas? There must also be within us a corresponding unchanging metaphysical component, the mind (*nous*).

Augustine of Hippo reworked this inner/outer pair of dualisms (Ideas/phenomena ↔ mind/body) into the specifically Christian form of the God/created universe ↔ soul/body. This transformation, however, produced a rather severe straitening of the intellect. In the Platonic world, anyone—having a mind—potentially can gain metaphysical insight. In the Christian version, however, the inner light by which we can read the book of Nature is bestowed upon us only through God's inscrutable grace. Moreover, Platonic Ideas are somewhat heterogeneous: there is the Idea of a table, the Idea of a feather, and so on, as well as the more difficult Idea of the Good. In contrast, the book of Nature reveals the singular if ultimately incommensurable will of God. The meaningfulness of the phenomenal world in its particularity collapses, and to turn one's will away from God and toward the distractions of the created universe is sinful delusion.

Renaissance thinkers were not quite so rigorous in their sense of the alienation of God from his handiwork, but the basic model persisted. Grace remained central but uncertain. Those who wrestled with doubts about the withdrawal of inner light confronted the “dark night of the soul.”

In this context, Descartes seems positively sunny. Through his own doubts he discovers one certain proposition that needs no special grace to support it: he thinks, therefore he is. Granted, the “he” who exists is a *res cogitans*—a thinking thing—while the rest of him—the *res extensa* (the thing with extended physical existence)—is only a reasonable conjecture. Luckily, however, any concept that reason presents to us clearly and distinctly must surely be true since it would be a slander against God to suggest that He “made us so imperfectly that by using reason rightly we nevertheless went wrong.”<sup>2</sup> Within this rationalist framework, exploration of the physical world becomes not only possible but in fact commendable, because “it may also be a great service to form a worthy judgment of God’s works, and possess that idea of the vast extent of the universe” which Descartes presents in his writings.<sup>3</sup> The phenomenal world becomes intelligible through the mediation of God-granted reason. However, in restoring the connection between the *res cogitans* and the *res extensa*, this divine mediation is central to experience, and Descartes’ program for free inquiry by the individual

---

<sup>2</sup> Rene Descartes, “The Visible World,” *Principles of Philosophy*, Part 3, in Elizabeth Anscombe and Peter Thomas Geach, eds., *Descartes Philosophical Writings* (Indianapolis: Bobbs-Merrill, 1954), p. 224.

mind therefore remains one more transmutation of Platonic dualism that shows the enduring power of the underlying form.

In the early empiricist tradition, skepticism arose to undermine Descartes' claims about the relationship between the *res cogitans* and the *res extensa* and about the authority of clear, distinct ideas in our understanding of the experiential world. David Hume in particular argued that the evidence of our senses allows us to draw no sure conclusion. What happens in our minds—our operations of inference—has no direct connection to what happens in the world. Immanuel Kant, following on Hume, argued that even confidence about what happens in the mind—and our privileged access to it—is unwarranted.

Kant's approach to knowledge of the world presented in his *Critique of Pure Reason* rewrites the Platonic phenomenal/noumenal distinction and transforms the noumenal into a pale ghost of its former self. It becomes a set of ineluctable categories within the organization of our experience of the phenomenal. Like Hume, Kant is uncertain about the status of causality, but he argues that we cannot conceive of experience of the world outside of the web of causal processes. Similarly, we cannot conceive of experience outside of space and time, but we have no right to draw ontological conclusions from this constraint. Finally, we cannot escape the relationship between a subject who experiences and objects that are the stuff of experience, but we have no right to assume their existence. Instead Kant treats a “transcendental” self that shapes experience into a coherent whole as a

---

<sup>3</sup> Descartes, Letter to Princess Elizabeth, 15 September 1645, in *Descartes Philosophical*

necessary postulate we simply have to live with but about whose existence we can remain neutral. Correspondingly, we have to postulate the existence of “things-in-themselves” in order to talk about the world of experience, but we cannot derive any further information from this merely formal assumption. We have an opaque dualism without either the certainty of a mind or the intervention of grace.

By sharply curtailing the claims of metaphysics (i.e., talk about the grand themes of God, Freedom and Necessity and about real minds knowing real objects) Kant makes room for the limited claims of empirical science. In the later *Critique of Judgment*, Kant moves his own account away from ontological system-building and into the realm of a philosophical anthropology that describes and explains the nature of human thinking without grounding its particular characteristics in necessity. In the *Critique of Judgment* Kant returns to the issue of the validity of the types of universal categories he proposed in the *Critique of Pure Reason* by asking where categories of any sort come from. Since all we have are singular sensations, how do we justify the grouping of particular “things” into larger, more general categories? Why did we even consider the possibility of these groupings in the first place? His answer is from the realm of philosophical psychology rather than from first principles: we experience moments of recognition of a coherence that oversteps all conceptual knowledge. We apprehend “These two things belong together,” and then on the basis of the intuition we attempt to understand the logic behind this judgment of coherence. This initial judgment based simply on the

correspondence of what appears in the world before us with our faculties for perceiving that world is an *aesthetic judgment*.<sup>4</sup>

Kant sees one basic, necessary assumption underlying these aesthetic judgments, namely, that nature as a whole must be coherent at least to our human abilities to perceive.<sup>5</sup> When humans make aesthetic judgments that prove viable when subjected to further inquiry, they in effect have revealed an aspect of nature's formal order. The quality which intuits these rules is genius. However, although all forms of knowledge rely on the progressive assimilation of the rules for order in nature revealed by genius, the realm of human endeavor that explores aesthetic judgments in their pure, preconceptual form is that of the fine arts.<sup>6</sup> While Kant's

---

<sup>4</sup> As Kant explains: "Now if in this comparison, a given representation unintentionally brings the imagination (the power of a priori intuitions) into harmony with understanding (the power of concepts), and this harmony arouses a feeling of pleasure, then the object must thereupon be regarded as purposeful for the reflective power of judgment. A judgment of this sort is an aesthetic judgment about the object's purposiveness; it is not based on any concept we have of the object, nor does it provide such a concept." Immanuel Kant, *Critique of Judgment*, trans. Werner S. Pluhar (Indianapolis: Hackett Publishing, 1987) p. 30.

<sup>5</sup> "In a critique of judgment, the part that deals with aesthetic judgment belongs to it essentially. For this power alone contains a principle that judgment lays completely a priori at the basis of its reflection on nature: the principle of a formal purposiveness of nature, in terms of its particular (empirical laws), for our cognitive power, without which principle the understanding could not find its way about in nature." Kant, *Critique of Judgment*, pp. 33-34.

<sup>6</sup> Kant concludes: "(1) Genius is a *talent* for producing something for which no determinate rule can be given, not a predisposition consisting of a skill for something that can be learned by following some rule or other; hence the foremost property of genius must be *originality*. (2) Since nonsense too can be original, the products of genius must also be models, i.e., they must be *exemplary*; hence, though they themselves do not arise through imitation, still they must serve others for this, i.e., as a standard or rule by which to judge. (3) Genius itself cannot describe or indicate scientifically how it brings about its products, and it is rather as *nature* that it gives the rule. That is why if an author owes a product to his genius, he himself does not know how he came by the ideas for it; nor is it

version of artistic genius was foundational for his epistemology, it was also highly circumscribed in its application as well as in its theory. For example, the empirical difference between products of genius and those of sheer random association was an after-the-fact judgment: did they reveal exemplary rules that could serve as guides for others? Moreover, the ontological status of those exemplary orderings remained radically uncertain: they work for us, given our capacity for experience, but we have no idea if they are *real*.

How European intellectual culture got around the impasse of Kant's critique of knowledge has added several more important strands to our current habits of thinking about subjects and objects, experience and meaning. Chief among these is G. W. F. Hegel's transformation of the contradictions encountered in the *Critique of Pure Reason* when Kant attempted to apply his categories to see if they could reveal substantive knowledge about the world. Hegel objected to Kant's refusal to say more about the self behind the categories or about the world which the self so remarkably synthesizes into coherence. He complained that Kant ignored the implications of his own arguments.<sup>7</sup> Hegel argues that the dynamics of Kant's

---

in his power [*Gewalt*] to devise such products at his pleasure, or by following a plan, and to communicate [his procedure] to others in precepts that would enable them to bring about like products. (Indeed, that is presumably why the word genius is derived from [Latin] *genius*, [which means] the guardian and guiding spirit that each person is given as his own at birth, and to whose inspiration [*Eingebung*] those original ideas are due.) (4) Nature, through genius, prescribes the rule not to science but to art, and this also only insofar as the art is fine art." Kant, *Critique of Judgment*, pp. 175-76.

<sup>7</sup> Hegel explains: "Certainly the forms of thinking should not be used without investigation, but this process of investigation is itself a process of cognition. So the activity of the forms of thinking and the critique of them, must be united within the process of cognition. The forms of thinking must be considered in and for themselves; they are the ob-ject and the activity of the ob-ject itself; they investigate themselves, [and]

formal place-holding postulates—i.e., the transcendental self and a world of things-in-themselves—must describe how the world really works. The self-contradictions Kant encountered in systematically applying each of two mutually exclusive attributes that derive from his categories—like freedom or necessity, or finitude or infinity of space and time—to the world of experience show not the failure of reason to describe the world but an actual dialectical relationship between the self and the world. Hegel describes this ontologizing of Kant’s critique as a shift from subjective to absolute idealism.

The post-Kantian Romantics followed a trajectory in their appropriation of Kant that was similar to Hegel’s but not quite so systematic, thorough, or dialectical. The self that guaranteed the unity of apperception was neither a mere transcendental postulate nor a self-alienated aspect of Absolute Spirit but ones very own concrete, particular self. However, this self, preserving Kant’s subject-object distinction, reverted to the older Cartesian dualism of a substantive mind as *res cogitans* (“thinking thing”) and the derived body as *res extensa* (“thing subject to extension [in time and space]”) even though poets continued to chafe under the division. Instead of God-given Reason mediating between the mind and the world of objects, genius became both the creator and the guarantor of the intelligibility of

---

they must determine their own limit and point out their own defects.” G. W. F. Hegel, *The Encyclopaedia Logic*, T. F. Geraets, W. A. Suchting, and H. S. Harris trans. (Indianapolis: Hackett, 1991), p.82.

the world.<sup>8</sup> Works of genius—still ultimately outside the claims of causality and still capable of creating regulative rules for experience—were real, indeed more real than mundane appearances. Correspondingly, the phenomenal realm remained meaningful, but it attained meaning not in itself but through the agency of the human imagination. The emancipation of the human self was heady stuff. It validated fiction and the Imagination and affirmed the new, self-sufficient individual of the nascent bourgeois society.

However, Western literary culture in particular has been paying for the hubris of throwing off the Kantian constraints ever since. Both knowledge of the world and the authenticity of the self are based on the ability of human genius to freely produce the rules by which we live. Maintaining the possibility of free creativity within the phenomenal realm, however, has proven too great a burden for human imagination to sustain. It has shifted and buckled in unexpected forms under the weight.

The failure of genius and imagination as a mode of knowing the world occurred at many levels.<sup>9</sup> The most compelling and immediate failure was experiential: the phenomenal realm proved singularly resistant to human desire. Shelley's "Prometheus Unbound" gave way to his "Triumph of Life." Poets and

---

<sup>8</sup> M. H. Abrams explores these transformations in *The Mirror and the Lamp: Romantic Theory and the Critical Tradition* (New York, Oxford University Press, 1953), his classic study of the imagination in English Romantic literary thought.

<sup>9</sup> The overview that follows is very general. Since, however, most of what I describe are truisms from the study of nineteenth century cultural history, so I do not elaborate. I freely admit that Romanticism is a very loose term and that Keats' Romanticism and



thinkers discovered that Nature, which Kant had circumspectly postulated to be purposive, proceeded by designs indifferent to human interests.<sup>10</sup> However, since epistemology was bound to genius (“Nature, through genius, prescribes the rule not to science but to art...”), this inhuman Nature turned opaque. The world of objects, being resistant to human will, became either malevolent (“red in tooth and claw”) or utterly alien. That the natural sciences were making significant discoveries over the course of the nineteenth century just added insult to injury: where Imagination failed to create rules for the world of phenomena, careful thought and observation succeeded in discovering them.

### From Individual to Collective Genius

The assault on the power of individual genius to mediate between the self and the world and to shape meaning within the phenomenal realm has taken, as well, a theoretical turn to explain the empirical failure. The attack has two fronts: the

---

Wordsworth’s, Byron’s, Schiller’s, or Hugo’s are all different. I believe, however, that the general account here captures central shared features.

<sup>10</sup> Kant’s understanding of the term “purpose” is a bit different from current usage:

Now insofar as the concept of an object also contains the basis for the object’s actuality, the concept is called the thing’s *purpose*, and the thing’s harmony with that character of things which is possible through purposes is called the *purposiveness* of its form. Accordingly, judgment’s principle concerning the form that things in nature have in terms of empirical laws in general is the *purposiveness of nature* in its diversity. In other words, through this concept we present nature as if an understanding contained the basis of the unity of what is diverse in nature’s empirical laws.

For Kant, then, the purposiveness of nature was the assertion that, from the human perspective, all the laws of nature and all the phenomena observable in nature have a coherent causal ground. Humans may not necessarily play an important role in this ground: we impute a coherence even if it is not of our making.

relationship of the individual to society and the status of language as a human artifact. In each case, the Romantic model for the human creation of meaning in the world of objects changed form rather than collapsing altogether.

Within the terms of the Romantic version of Imagination, an individual who is not autonomous cannot reasonably claim to create anything and is instead just the medium through which large-scale processes unfold. Either one is a Subject who proceeds by laws of Freedom, or one is an object trapped within a world of Necessity. Twentieth century thought has given the idea of autonomy a rather severe beating, and coherence of meaning mediated by genius has suffered with it. Anthropology, sociology, and psychology trace the social factors that structure the ego: they see it as an effect rather than a cause. Some versions of philosophy view the self as sheer humbug. Others, more sympathetic to claims of meaningfulness in literature, depict the individual as deeply shaped by historical context. Since freedom and necessity are the terms defining the choices in the individual's engagement with the world, it would seem that the more individuals become defined by the causal patterns in which they develop, the more they fall from freedom into necessity. Nonetheless, a curious twist arises. At one extreme end of determinism is one variation on social constructionism wherein the causal determination of meaning becomes specifically social and therefore once again an entirely human affair. In this model Nature disappears entirely, with the result that society's *collective* genius ends up giving the rule to humankind. Nature, banished from human self-construction, becomes unspeakable, and objects withdraw into silence.

Through all these changes, the continuity of terms from Romanticism to post-Romantic movements and to contemporary post-modernism is very strong: the human mind creates what meaning it finds in its relation to the phenomenal world, even if the locus of this creation gradually is displaced from the individual to the collective. This now collective human mind, moreover, somehow remains the Cartesian *res cogitans* floating free from the *res extensa* of the brains that house it. Consider, for example, how the feminist scholar Judith Butler in *Bodies that Matter* attempts to assimilate the body and materiality into the shared social world of language:

The body posited as prior to the sign, is always *posited* or *signifies* as *prior*. This signification produces as an *effect* of its own procedure the very body that it nevertheless and simultaneously claims to discover as that which *precedes* its own action.<sup>11</sup> If the body signified as prior to signification is an effect of signification, then the mimetic or representational status of language, which claims that signs follow bodies as their necessary mirror, is not mimetic at all. On the contrary, it is productive, constitutive, one might even argue *performative*, inasmuch as this signifying act delimits and contours the body that it then claims to find prior to any and all signification.

This is not to say that the materiality of bodies is simply and only a linguistic effect which is reducible to a set of signifiers. Such a distinction overlooks the materiality of the signifier itself. Such an account also fails to understand materiality as that which is bound up with signification from the start; to think through the indissolubility of materiality and signification is no easy matter. To posit by way of language a materiality outside of language is still to posit that materiality, and the materiality so posited will retain that positing as its constitutive condition. Derrida negotiates the question of matter's radical alterity with the following remark: "I am not even sure that there can be a 'concept' of an absolute exterior." To have the concept of matter is to lose the exteriority that the concept is supposed to secure. Can language simply refer to materiality,

---

<sup>11</sup> The similarity to Hegel's move in rescuing the self from Kant's negative, transcendental analysis is striking. See note 7.

or is language also the very condition under which materiality may be said to appear?<sup>12</sup>

The argument is a good example of contemporary theoretical analysis of the problem of concepts in general, and it suggests the great distance that must be overcome between current theory and any possible biological account of phenomenal experience. Butler's argument is entirely valid as far as it goes. Yet the choices she offers are caught up in the Cartesian dichotomy of *res cogitans* and *res extensa*: matter is supposed to be "absolutely exterior" to language. Just as the mind as *res cogitans* was part of Descartes' defense of the integrity and indubitable existence of the individuated self, Butler offers a defense of the autonomous self-creation of a collective social world of language. The problem is that this approach attempts to give language as such the transcendental status of an a priori condition for the possibility of experience. This approach does not work. Language lacks the austere inscrutability of a transcendental category: we know far too much about its empirical logic.

### **The Linguistic Structuring of Meaning and the Biological Turn**

The vision of a human community that collectively determines the structure of the world of experience through its creation of language has proven a powerful tool of social critique. Contextualizing this mode of analysis within a biological account of the structures and processes that shape language does not diminish the

---

<sup>12</sup> Judith Butler, *Bodies that Matter* (New York and London: Routledge, 1993), pp. 30-31.

incisiveness of the insights offered by the “linguistic turn” in cultural criticism but does recast the theoretical models by providing a microstructure for meaning.

Butler suggests that language is “the very condition under which materiality may be said to appear.” This view of language as the horizon that determines the possibilities of phenomenal experience has grown out of 20<sup>th</sup> Century linguistic theories that provided a general model for the organization of sign (semiotic) systems. Semiotic theory in turn has become increasingly sophisticated and extended its reach to encompass all forms of cultural production, but its development has remained largely separate from the advances in the empirical sciences that explore the intimately related realms of perception, cognition, and language.

The idea of living within the horizon of linguistic practice evolved from an important shift early in the 20<sup>th</sup> century in the understanding of how language works. That the relationship between words and what they signify—WATER as “water,” “agua,” “eau,” or 水 — was basically arbitrary had been known for a long time. Within this arbitrariness, the binding of meanings to sounds was seen largely as a genealogical process relying on strong historical continuities. The number of synonyms shared by Greek and Sanskrit, for example, offered stunning proof of this model of inheritance. Ferdinand de Saussure offered an important new perspective when he separated diachronic (“across time”) from synchronic (“at the same time”) processes for the attribution of meanings to sounds. He argued for “meaning through difference” within a synchronic structure. That is, both of the

significant components of a word—the signifying sound and the content of the signified concept—are grounded in structures of such components that define one another through mutual differentiation. Each language has a subset of all possible sound elements (phonemes) which—when absent or present—mark different words. In the language, the distinctions not in the signifying subset simply will not matter. Japanese, for example, does not distinguish between the “l” and “r” sounds, while English does not separate the “d” and the aspirated “dh” used in languages like Sanskrit. Similarly, though not so obviously, there are patterns of mutual differentiation that define domains of possible meanings for signified concepts. Consider, for example, the sorts of images associated with “shrub,” “bush,” and “tree.” Each category has its own distinctive prototypical examples, but each also blurs into the next at the boundary. Different languages divide vegetation along different lines with either greater or fewer distinct categories. Structural linguistics argues that in principle one should be able to derive a set of binary attributes (i.e. only allow the values “yes/present” or “no/absent”) to explain the logic of differentiation within the language. In some languages, particular botanical features would matter, and in some they would be irrelevant to the classification scheme.

The fact that the *content* of linguistic representations is a socially determined structure gives rise to the possibility that our entire world of meaning is self-contained and self-determined. This possibility, however, rests on the idea that the mechanisms for the production of language are free and unconstrained by the mediation of underlying biological processes. Once again, though, the mind as *res*

*extensa*—a thing within time, space, and causality—intrudes. The central problem in the material mediation of meaning is where the categories of differentiation come from. At the level of phonology—the signifying sounds—at least a partial answer is fairly clear. The physical means of production—i.e. the features of the mouth, nose, and throat—create the basic components that give structure to the system. Why any particular language has the sound structure it does, however, is not so obvious.

The origins of the categories that differentiate the conceptual content of words are even less clear. Post-structuralist critics have long rejected as naive the idea that these categorical distinctions are simply random and arbitrary. Given the sorts of distinctions that most concern these critics (racism, sexism, class biases), they do not look to biology, which is seen as part of the problem. Instead, they argue that the conceptual distinctions that structure what can be represented in a language are social constructions that shape the distribution of power within a culture. In the post-structuralist view, mutually determining categories (like center/periphery or male/female) slide into relationships of inequality between initially coequal terms. This argument holds out the possibility of dissolving the power relations based on these reified inequalities and offers a vision of the free self-determination of meaning within human society.

This social constructionist account of the origins of conceptual categories in language is, for the most part, an explanation of the empirical organization of the mind. It tells us about the social processes by which we acquire the distinctions that matter in our experience of ourselves and of the world. It has proven a powerful

theory to account for social inequalities and has provided a basis for action to redress historical wrongs. It is, however, not a complete account. As a causal explanation, the social-constructionist approach to the logic of concept formation is complemented by the systematic empirical work of developmental and cognitive psychologists and neuroscientists. The cultural critic's reliance on large-scale social structures of power to explain the emergence of microstructures of conceptual organization are inadequate by themselves: in the end, biological structures rely on biological mechanisms. While radically free human self-determination will fail in the face of the causal structures of biological ordering, a central theme of this book is that much of the argument will remain intact. Fundamental structuring processes may be biological in nature, but layer after layer of higher order organization depends upon responses to the particularity of environmental conditions, and much of the human environment is humanly structured. As Vygotsky argued long ago, much of who we are and what we think is a reassimilation of the human manipulation of the world.

### **The Project of this Book**

The various versions of the linguistic model that prevail in contemporary humanistic discussions implicitly address the large issues of Freedom and Necessity, Genius and Nature, Self and World, and Subject and Object. These questions form an important but largely hidden component of the way in which we as a society think about the nature of meaning. I therefore seek to chart a rough path between microstructures of neuronal processes and the organization of experience mediated



by the large-scale structures of linguistic (or more generally semiotic) meaning. If we wish to understand the fine structure of semiotics, we need to learn about its grounding in neural organization and behavior. Conversely, because of the mediation of ever-higher order systems of representation within the brain, neuroscience by itself cannot hope to give an adequate account of the logic of experience. Humanistic expertise in the complexities of meaning in symbolic experience is an essential part of the story of our biological engagement with the world.

The biological account of meaning, however, does not present a resolution of the issues of knowledge explored in this chapter. There is a strong version of the social-constructionist argument which proposes that empirical explanations of how our semiotic structures arise—including the neuroscientific account provisionally offered in this book—cannot hope to securely ground our knowledge of the world. Philosophers like Donald Davidson and Richard Rorty have declared the Western epistemological project of connecting subjects with objects to have reached a hopeless impasse. Their arguments are a form of Kantian skepticism that agrees with Hegel in finding the transcendental ego and the object-in-itself merely formal postulates that yield no positive information about experience. They suggest giving up such categories and admitting that truth is a human affair and will have to be whatever we deem to be true through our best procedures. Rorty, in explaining Davidson's views argues:

We can tell an adequate story about the progress of human inquiry (in all spheres—logic and ethics as well as physics) by describing the continual

reweaving of systems of belief and desire. This reweaving is made necessary by the acquisition of new beliefs and desires—e.g. the sort that are caused to occur in humans by such events in the World as the opening of doors. (And ... by the invention of “successful” metaphors.) We do not need to raise the question ... of whether there are things in the world which make algebraic or moral truths, or aesthetic judgments, true. For although there are causes of the acquisition of beliefs, and reasons for the retention or change of beliefs, there are no causes for the *truth* of belief.<sup>13</sup>

Whether intended or not, the model for this argument seems to be the account of aesthetic judgments that ground epistemology in Kant’s framework. Just as there is no sure ground for the categories created by genius but only *post hoc* judgments about their exemplary nature, so too there is no back door leading to “things in the world” that might ground belief but only the empirical fact of “successful” metaphors. Moreover, since causality is a category that remains uncertain for Kant—a necessary part of human experience but ontologically inscrutable—for Kant, as for Rorty, there can be no “causes for the truth of belief.” Where Rorty’s anti-realism radically differs from Kant’s account of judgments is in the assertion that we “do not need to raise the question ... of whether there are things in the world which make algebraic or moral truths, or aesthetic judgments, true.” This claim strikes me as a form of whistling past the graveyard that—whatever its theoretical merits—is empirically false. Confronting the limits of our representational capacities is a powerful aspect of human culture, and indeed the Greeks had a word for the refusal to acknowledge the limits of human understanding: hubris.

---

<sup>13</sup> Rorty’s emphasis. Richard Rorty, “Non-Reductive Physicalism” in *Objectivity*,

In this book I seek to avoid three forms of hubris. The first is asserting that neuroscience finally will tell us how we know the world. It cannot. The anti-realist critique, like Kant's analysis of judgments, forcefully reminds us that neuroscience remains within the constraints of all empirical knowledge. The second type of hubris is the slightly more complex form offered by Rorty, that of arguing that what lies beyond the neuroscientific account does not matter. This book explores how we, as specifically *biological* creatures, experience the world. The neurobiologically conceived self is central to this story: it is the nexus of systems that synthesizes, judges, and acts. Our understanding of this neuronal self is the site of a third, and greater temptation to hubris. The dominant models of the self current in our culture—the Christian, Cartesian, or linguistic—all are shaped by a dualism in which the self maintains its integrity as a soul or *res cogitans* in contradistinction to the body. The biological version of the self—which emerges out of interactions between the environment and the developing cortical systems—seems to be a self-contained empirical account. Yet it is not so. The neurobiological model also inevitably lives within a dualism internal to the self that no empirical account can efface. The transcendental subject's shadowy tracework—the never quite visible but ever-present formal ordering of the world beyond the possibility of self-knowledge—is woven into experience and marks the radical incompleteness of any account of the empirical self. If we quite reasonably point to the limitations of a snail's understanding of the world, we must also acknowledge that we too can know only what and how we are equipped to know, and there is no compelling

reason to think our encounter with the world is much more adequate than the snail's. To believe otherwise would be hubris indeed.

The philosophical tradition—and Kant in particular—has explored the formal logic of category-formation and the *a priori* structuring of experience. This tradition offers us a framework within which to investigate a neuroscientific account of experience and selfhood that inevitably confronts similar (but not identical) issues in a very concrete, empirical manner. Building that framework has been the goal of this chapter. We know the issues at stake: the nature of the phenomenal self, and the logic of our engagement with the world of experience. We have some sense, based on an overview of the epistemological tradition, of what can and cannot be expected from the neuroscientific account. The rest of the book explores the current neuroscientific understanding of perception, cognition, memory and emotion that underlie our experience of the self and the world. In order to take this biological account seriously—not as a metaphor, but as the actual systems that *are* experience as we know it—the book is tightly focused on the neural mechanisms, cortical structures, and the developmental neurobiology that leads to the emergence of the self that we know from common experience.

The order of presentation is straight-forward. The second chapter, “Neurons, Matrices, and Models,” introduces the basic features of contemporary neural network models. Artificial neural networks are simulations of massively parallel arrays of interconnected simple processors that are inspired by the yet more massively parallel and densely interconnected neuronal organization of the brain. Important properties of neural networks emerge out of system-wide dynamic

interactions that are modeled through the powerful computational techniques of matrix mathematics. To begin to think about experience in biological terms—to move from metaphors to mechanisms—requires a clear general sense of what is involved in matrix mathematics. Therefore, the chapter also presents an overview of matrices before considering the ways matrix models emulate the functioning of neural networks in the brain.

The third chapter, “Connecting,” gives an overview of the brain’s neural systems. It begins with a discussion of neurons and synapses, then briefly outlines current understanding of the functional organization of the brain. It next turns to a more detailed account of the visual system from the retina through the thalamus and visual cortex and into the higher cortical processing areas. The chapter concludes with a discussion of memory systems and the top-down and bottom-up interactions between the sensory cortical area and those that process much more highly abstracted representations of objects and events.

The fourth chapter, “What Matters,” continues to explore issues of the cortical structuring of representation. The focus turns in particular to the mechanisms by which the brain comes to recognize what matters to the organism and the impact of those mechanisms (a.k.a. affect or emotion) on the structuring of knowledge of the world. Since fear has been widely studied, it serves as the paradigmatic case. The chapter argues in particular that, given the nature of neural networks and what we know of the memory process in the brain, the constraints of the body as represented by affective responses are central to the structuring of memory and meaning. The chapter then concludes by drawing comparisons first

between this proposed logic of affective organization and approaches to designing autonomous robots and then between the neurobiological model of affect and Buddhist, Chinese, and Heideggerian views of how we engage objects in the world.

The fifth and final chapter, “The Human World,” builds upon the role of affect in organizing meaning by exploring problems of early development. Here emphasis moves from the micro-level analysis of neuroscience to the transitional perspective of developmental neuroscience and then to various approaches in developmental psychology and psychoanalysis. This chapter presents an emerging synthesis of disciplines that traces the development of the neuronal self from the inchoate structures available within the early cortex through the successive stages of internalization of external structures, abstraction, and maturation. This biological self is one distinctive aspect of a set of neural systems inside a brain inside a body. Yet I hope that by the end of this final chapter the reader will see that this neuronal self encounters a world of great depth and variety and, in acting, draws upon a rich inner life of memories, meanings, desires and hopes made possible through the neural structures of the brain.

Every topic covered in this book—from the varieties of Hebbian learning algorithms to the biophysics of synaptic spiking, the functions of the hippocampus, the role of the amygdala, or the interactions between attachment and stress—has a great depth of scholarship that I can present only in overview. Most of these topics, moreover, are subjects of intense research, with many uncertainties and not a few contradictory findings. My surveys of these fields therefore inevitably are partial

and stress the particular issues that are important to develop the themes of this book. In order to suggest the nature of the research literature from which I have drawn my arguments and to allow interested readers to pursue these topics on their own, I have included an extensive bibliography of sources I have found useful. I hope that readers may see through the limitations of my presentation to recognize the importance of neuroscience to the broader inquiry into human experience and to use these sources to join in what surely will be one of the most important intellectual endeavors of the Twenty-first Century.