

ON THE *NEURO*-TURN IN THE HUMANITIES

*Naturalism, Hyper-Empiricism, and Understanding**

WILL SAMSON

Since the 1970s or so, and right up to the present day, a lot of work has been done to bridge humanistic meaning-making systems with those of the sciences, and of neuroscience in particular. From the late Francisco Varela's seminal work on phenomenology and neuroscience in the 1990s, to the work of Antonio Damasio and Catherine Malabou more recently, and including literary variations on the same theme, such as Ian McEwan's 2005 novel, *Saturday*,¹ we see that scholars and writers in the humanities are attempting to incorporate neuroscientific insights into their literary, political, philosophical, and theoretical work. It is no longer only analytic philosophers interesting themselves in neuroscience, but also continental philosophers and artists. Interestingly, where one might expect neuroscientists to have recourse to analytic philosophy of science in trying to explain their field, in the cases of Damasio and Varela, we find neuroscientists going to continental philosophy in order to contextualize and expand on their scientific work.

I became acutely aware of the strength of the neuroscientific trend in humanities research not very long ago when I was reading Varela and noticed that a number of people around me were looking into various similar areas: an acquaintance was reading a book on neuro-aesthetics and their applications to literature; a friend was reading a book on cognitive science and performance theory; another friend recommended a book on neuroscience and political theory to me; a number of friends, acquaintances, and strangers were writing theses or dissertations on philosophy, using thinkers who had one foot in the humanities and one in neuroscience. It seemed to me that there was something driving humanities scholars towards the neuro-humanities.

* The following paper and response were originally delivered on 28 November 2014 as part of the Theory Sessions at the Centre for the Study of Theory and Criticism in London, Ontario.

1 Ian McEwan, *Saturday* (New York: Anchor, 2006).

I asked myself, why neuroscience in particular? I'm aware of other lines of scientific thought that have been appropriated by the humanities over the years—I recall a collection of Darwinian literary criticism titled *Madame Bovary's Ovaries* which seems to be a particularly odious example of the trend—but none of these has had the seductive power of what can increasingly be characterised as the neuroscientific *turn* in the humanities. While Darwinian literary criticism has had some popularity, it seems to be on the decline, as indicated by the collapse of its dedicated publishing platform, *The Evolutionary Review*, a journal launched in 2010 which ceased publishing in 2013. Moreover, Darwinian literary studies has met with significantly stronger resistance than the *neuro-* turn, with much of the attendant literature on the topic being quite critical of it.² I do not, however, intend to argue about why humanities scholars seem to be gravitating towards neuroscientific ideas in particular; it is possible that the uptake in this kind of research and production has as much to do with financial concerns as it does with more properly methodological or intellectual concerns. Rather, what I wish to interrogate is how to have a valid and productive collaboration between the humanities and neuroscience. In order to effect this project, I am going to evaluate specific aspects of the thought of Catherine Malabou, Francisco Varela, and Antonio Damasio—the scope of this enquiry is thus limited, but I see these thinkers as being indicative of larger trends. Then I will propose some means of fruitful interaction between continental philosophy and neuroscience that do not fall into the same traps as Malabou, Varela and Damasio. My basic thesis is that the ways in which neuro-humanists currently proceed are often methodologically objectionable, and that there are better ways for humanists to collaborate with the sciences in general and neuroscience in particular.

Naturalism and the Humanities

One aspect of why the *neuro-* turn might be appealing to a lay reader is that it invokes naturalism in contexts that are sometimes seen as incred-

² See, e.g., Stephen Pinker, review of *The Literary Animal: Evolution and the Nature of Narrative*, eds. Jonathan Gottschall and David Sloan Wilson, *Philosophy and Literature* 31 no. 1 (April 2007): 162-78; Eugene Goodheart, *Darwinian Misadventures in the Humanities* (New Brunswick, NJ: Transaction Publishing, 2007); Jonathan Kramnick, “Against Literary Darwinism,” *Critical Inquiry* 37 no. 2 (Winter 2011): 312-47.

ibly subjective. Naturalism proceeds by exposing causal relations in such a way as to disallow any explanation that falls outside of the laws of nature narrowly construed. To make the claim that a naturalist approach can be turned towards the analysis of phenomena that are traditionally seen as being outside of such relations is a seductive claim, regardless of how difficult it might be to substantiate. Put differently, naturalism did not firmly take root in the humanities for various reasons,³ and the *neuro-* turn seeks to ground humanistic research in a naturalistic paradigm, one based on neuroscientific insights. In such an account, neuroscience, as that realm in which it is now becoming possible to treat objectively questions which philosophers have hitherto only been able to speculate about, becomes first philosophy.

The relationship between the neuro-humanities and naturalism is, perhaps surprisingly, fraught. One problem is that scholars in the humanities are generally not trained to think in the methodological terms of the sciences, even though it would behoove thinkers engaging in the neuro-humanities to have a deep enough appreciation of neuroscience to incorporate elements of its disciplinary framework. A parallel problem is that scientists who take on the task of relating their work to the humanities should have a better idea of how their naturalistic research would interact with the kinds of value-objects to which they endeavour to relate it. In the work of Catherine Malabou, we find some form of the first problem: there is something else at stake in Malabou's project making her take on only superficial elements of neuroscience and read them metaphorically rather than substantively.

In what is perhaps Malabou's best work on the topic, *What Should We Do With Our Brain?*, she argues that the notion of *neuroplasticity* operative in the discourse of neuroscience is cognate with the notion of *flexibility* operative in neo-liberal capitalism. As she sketches in her conclusion to the book, the answer to the titular question, "what should we do with our brain?" is that we should resist:

to ask "What should we do with our brain?" is above all to visualize the possibility of saying no to an afflicting economic,

³ While a closer look at the origins of the disciplinary distinctions between the humanities and the sciences would be out of place here, suffice it to say that the objects of study of the humanities seemed, for a very long time, to be safely locked away in intra-cranial darkness and thus insulated from the tools of the sciences, but open to humanistic methods such as introspection and textual hermeneutics.

political, and mediatic culture that celebrates only the triumph of flexibility, blessing obedient individuals who have no greater merit than that of knowing how to bow their heads with a smile.⁴

This answer, regardless of whether we agree with it, approaches neuroscientific data from an idiosyncratic point of view. Indeed, it seems that the conclusion here is completely distinct from the discussion of the notion of plasticity at play in the neurosciences, and that a similar prescription—refuse to be flexible, *resist!*—could be made without reference to neuroscience at all. What she is actually engaged in is not a synthesis of neuroscience and the humanities, but an attempt to ideologically correct the discourse of the sciences. In *What Should We Do With Our Brain?* she points to multiple pernicious norms hidden in neuroscience’s descriptive statements—reading plasticity as flexibility in line with the *modus operandi* of neo-liberal capitalism is just one of them. This is the deconstructive game—her commitments, be they to materialism, naturalism, Hegelianism, or to deconstruction itself, matter little. What is important for her project in this instance is to construct a philosophical non-space from which to point out, criticize, and correct the normative claims of other discourses.

However, her stated goal at the beginning of the book was to “awaken a consciousness of [the constitutive historicity of the brain].”⁵ Given that the conclusion she arrives at is in no discernible way connected to this goal, she almost certainly fails. Yet what is the value of awakening such a consciousness? A thorough, sober, and informed discussion of how knowledge of the brain might change the way we think and act would be beneficial. Certainly, Malabou places emphasis on the historicity of the brain, but in such a way that the import of the brain in particular as a specific organ remains unclear. Malabou is open to the same criticism Alan Sokal levelled at continental philosophers a generation ago. He accused a number of continental philosophers of invoking scientific discourse in order to lend their claims an air of credibility that they do not deserve on their own. Malabou’s naturalism, then, seems to be opportunistic if not false—a fashionable position to take, which allows her to position herself better rhetorically, but which is

4 Catherine Malabou, *What Should We Do With Our Brain?* (New York: Fordham UP, 2008), 79.

5 *Ibid.*, 2.

neither necessary nor relevant to the political conclusions she wishes to draw.

Malabou's text, then, can serve as something of a warning. *What Should We Do With Our Brain?*² gives us a salient example of what happens when we conflate categories. The title and introduction of the book imply a naturalistic focus—reawakening a consciousness of the constitutive historicity of the brain is a project that should take up the naturalist or materialist insights of the sciences in order to increase the humanistic understanding of the formation of the self, of self-fashioning. On the other hand, much of the rest of the book, for all its lip-service to neuroscience and its invocations of terminology and relevant texts, has little to do with this initially announced project, and instead has recourse to textual hermeneutics that seek to unveil the hidden normative standards of the discourse. Theory, like most interdisciplinary work, allows us to approach a discipline from outside of its own self-understanding—but this can have a Frankensteinian effect. There seems to be a tear, haphazardly stitched over, between Malabou's naturalistic posture and her humanistic destination; one, moreover, that she does not adequately bridge. If one is correct in imputing a certain form of naturalism to Malabou, it is confined to the notion that there is no permissible ontological dualism, thus that nature and culture are logically continuous spaces. If, as I think, this insight is fundamentally correct, the question then becomes one of approach: how do we talk about phenomena in the humanities if they are essentially of the same stuff as the objects of the natural sciences? While their shared grounding in the same material stratum allows us to put these discourses into communication, Malabou leaves us fundamentally ill-equipped to actually stage such a conversation.

Hyper-Empiricism: the reduction to material as epistemology

In making her way from a bastard naturalism to a critique of the ideological underpinnings of scientific claims, Malabou brings into relief a tension between norm and nature that discomfits some aspects of the naturalistic discourse. In fact, what can come out of it, if we read her charitably, is a good reproach of what I will call hyper-empiricism—by no means a new concept, but something that requires some attention in this context. From at least Hume, we see a hypostatization of the “copy thesis” of empiricism, namely, that all ideas originate in sense impressions.

Today, the extension of this thesis culminates in statements from scientists and popularisers of science, such as those most recently made about the value of philosophy by Lawrence Krauss, Neil deGrasse Tyson, and Stephen Hawking—all of whom claim that philosophy is no longer the queen of the sciences, nor the handmaid to the sciences, nor even the humble under-gardener clearing away the brush, but a dead discipline that adds nothing of value. This hyper-empiricism comes out of an extension of the copy thesis when that thesis is read as grounding the objectivity of scientific discourse in such a way that insulates it from both the critiques of the tribunal of reason and from ideological contamination. Essentially, the statements of Krauss *et al.* emerge from the mindset that claims that because the objects of philosophy and the objects of natural science are of the same world, the most fundamental approach—the natural scientific approach—will be adequate to the task of describing the essential components of them both. In this account, the scientific worldview is sufficiently advanced to have done with philosophy’s mystery-mongering. But this is precisely what Malabou’s text shows not to be the case: scientific discourses are no more insulated from ideological contamination than any other—they are simply better positioned, rhetorically, to hide their normative claims and biases.

I term hyper-empiricism any form of empiricism that lays claim to a monopoly on explanation. Opposed to hyper-empiricism might be any number of forms of epistemic pluralism. Hyper-empiricism is seductive (and prevalent in the neuroscientific turn in the humanities) because it levels down the epistemic commitments of the enquiry at hand by displacing the question of how best to understand a given phenomenon onto what physical processes give rise to said phenomenon. The movement is subtle, seductive, and wrong, insofar as it elides the differences between ontology and epistemology, in short, turning a question about understanding into one of composition. Generally, we can see this in particularly facile attempts to reduce phenomenal consciousness or even self-consciousness to its cortical substrates—at the end of the day, says the reductionist, what we are dealing with is made of neurons, or of atoms, or of quarks or of strings. Ultimately, we are dealing only with these substrates. But there is a lot built into this “ultimately” and a lot of places this “ultimately” can lead us, depending on how far the reductionist rabbit has dug its hole. If we want an understanding about how the brain works, it is important to look to neuroscience.

Likewise, if we are looking for a full picture, it is important to see when neuroscientific insights might have some impact on higher-level processes. But to reduce these higher-level processes to their neural correlates is to commit a fallacy. There are a number of things that get left out when the jump is made from the phenomenal manifestation of a thing to the cortical substrates that underlie that manifestation—at the very least, those who would like to claim that the two are in some sense identical would have to concede that the latter misses out on the “what-it’s-like-ness” of the former.⁶

At this stage, there does not seem to be a good theoretical rationale guiding us with regards to the proper level to which to reduce things, and when doing so might be appropriate, which leaves only the practical justification: it works. “It works” can mean a lot of different things. To a practicing neuroscientist, it makes sense to focus on the neural level of analysis, since theorizing about quarks or strings is unlikely to result in localizing a brain function or allowing for the physical intervention on and manipulation of a function that the brain controls. The level of analysis makes sense given the practical ends to which a research project is oriented—it is not necessarily that the neuronal level is somehow ontologically or epistemically privileged, only that it becomes effective to think at that level *given some goal*. But does it make sense for humanistic research to yoke itself to that same level of analysis? Further, when we engage in neurohumanities, are we explaining the given phenomenon or explaining it *away*? It seems unlikely that the level of analysis appropriate to the study of literature, metaphysics, or politics would be limited to the neural or even fundamentally ground itself at that level of analysis. Certainly, an investigation of neural structures might be *pertinent*, but under the aspect of explanatory monism one might find it incomplete.

With this in mind, we turn to a book edited by an interdisciplinary team of scholars including Francisco Varela, a neurobiologist; Jean Petitot, a mathematician; Jean-Michel Roy, a philosopher; and Bernard Pachoud, a psychoanalyst. *Naturalizing Phenomenology* came out in 1999 to much fanfare, bringing together a diverse group of scholars in continental philosophy

6 See Thomas Nagel, “What Is It Like to Be a Bat?,” *The Philosophical Review* 83 no. 4 (October 1974): 435-50. The notion of “what-it’s-like-ness” (though not the phrase itself) stems from Nagel’s essay, but has generally been taken up as a means of highlighting the first-person aspect of consciousness. In short, if there is a mental state, there is something that it is like to be in that mental state.

and cognitive science in order to reconcile naturalism with phenomenology, which has seen the dangers of a naturalistic approach to consciousness since its Husserlian origins. The way in which the project attempts this reconciliation is obvious from the title—the point is to naturalize phenomenology, not to phenomenologize naturalism. Taking neuroscience as a kind of first philosophy, the editors of the book nonetheless claim that the dismissal of the first-person point of view, and of the phenomenological and subjective dimensions of cognition, constitutes a significant lacuna in the project of cognitive science, despite its claims to be the first truly scientific approach to consciousness. In the editors' view, the refined accounts of consciousness provided by Husserlian phenomenology cannot be ignored by cognitive science, and might enable thinkers to bridge or close the explanatory gap between scientific (third-person) accounts of the neurophysiological register of experience and the phenomenal (first-person) experience of consciousness. A renewed focus must be placed, not on what is going on inside the black box of consciousness, but on what is going on *for* the black box.

The problem with this work, which is the capstone and continuation of Varela's project of neurophenomenology, is that naturalizing Husserlian phenomenology is difficult: the way of going about it is not obvious. Even among the editors, the definitions at play and the solutions undertaken are so alien to each other that, despite their mutual affinities, it is difficult to say to what extent they agree. The impetus for naturalizing a theory of consciousness and subjectivity that has been so successful at describing and analyzing the phenomenal level of consciousness is clear—one would like to have all effective explanations exist in the same logical space, or at least not contradict one another. The dangers of such a project are equally clear: often, putting things in logically continuous space is much easier when we simply reduce one level of analysis to another, which, in this case, would bring us back to the problems of explanatory monism and the explanatory gap.

The editors of *Naturalizing Phenomenology* are mistaken about the nature of Husserl's rejection of naturalism, and their misunderstanding is enlightening with respect to the question of hyper-empiricism and explanatory monism in the *neuro-* turn. In their 80-page introduction to the volume, they defend their project by claiming that Husserl's rejection of naturalism regarding consciousness was based on the contingent limitations of scientific (mathematical) formalization at the time Husserl was writing. In their

view, Husserl's antinaturalism "is the result of the scientific limitations of his day, limitations Husserl thought impossible to overcome, making it thereby also impossible for the physico-mathematical sciences to offer a scientific reconstruction of the phenomenality of the surrounding world."⁷ This defence, and the elaboration of a minimal theory of interaction between the sciences and phenomenology that follows from it, misses a crucial aspect of Husserl's methodological antinaturalism. Husserl's strong negative reaction to naturalism was not based on the impossibility of mathematizing phenomenal descriptions, and therefore of integrating them into the general framework of the natural sciences,⁸ though he did hold this to be impossible. Rather, Husserl's antinaturalism was motivated by a transcendental objective, namely the constitution of meaning in consciousness and its directedness towards a world. It is quite beside the point to claim that one can mathematize and formalise a phenomenological description and thereby view consciousness as another (psycho-physical) object in the world—what is important is that doing so brings one no closer to determining how the constitution of sense works, how meaning is bestowed, what value-orientation is enacted at the level of experience, and so on. Certainly, in any intentional act, there are neurophysiological correlates that are produced, but what is manifested in the act is quite different from these correlates or substrates.

While Husserl was invested in a methodological antinaturalism, it was also his opinion that it is sufficient to effect an attitudinal shift in order to make phenomenological descriptions relevant to psychology (that is, empirical or functional psychology), or to the sciences that concern themselves with consciousness in general. Such an attitudinal shift is fraught with the danger of transforming phenomenology into descriptive psychology should one forget that phenomenology is first and foremost a science of essences and not of the empirical facts of an individual conscious experience.⁹ Making such a shift, however, involves a certain kind of methodological or epistemological pluralism, one more in line with Husserl's project than the project of *Naturalizing Phenomenology*. Varela *et al.*, despite their best efforts, ignore the properly philosophical reasons that Husserl avoided a naturalistic

7 Francisco Varela *et al.*, *Naturalizing Phenomenology* (Stanford CA: Stanford UP, 1999), 40.

8 *Ibid.*, 42.

9 Edmund Husserl, *Introduction to Logic and Theory of Knowledge: Lectures 1906/7*, trans. Claire Ortiz Hill (Dordrecht: Springer, 2008), 381.

standpoint, and, to their discredit, elide the differences between epistemology and a positive scientific investigation, differences that were important to Husserl's project.¹⁰ The claim of *Naturalizing Phenomenology*, that phenomenology can be subsumed into a naturalist discourse (specifically that of cognitive neuroscience), goes against the very nature of the Husserlian project, for reasons that should be of interest to any scholar in the humanities trying to bridge the gap between third- and first-person descriptions: one need not reduce the phenomenal to the neural in order to understand it. In fact, and here is where I think the fruitful moment of interaction between phenomenology and neuroscience can take place, it seems more likely that a proper understanding of the internal constitution of the phenomena and how they appear—the internal organization of that which is given phenomenally, highlighted by phenomenology—would be the necessary basis for a neuroscientific analysis of the cerebral and somatic underpinnings of such a phenomenon, or at least would provide valuable material for the interpretation of data. To claim otherwise is to risk falling into the same trap that Varela *et al.* do in eliding the difference between the empirical genesis of a subjective phenomenon and the validity and internal structure of a mental act.

A reductive point of view, *à la* hyper-empiricism or explanatory monism, leads us to an untenable position that sees the empirical processes through which the subject is engendered as the same processes to which the subject can be reduced. Rather than liquidating the subject entirely, making it into a mere objective being in the world, the neuronal correlates of experience provide functional information about the processes underlying consciousness. Phenomenology provides a way of understanding these physical explanations with reference to the lifeworld, contextualizing them along non-reductive lines, and producing satisfying *clarifications* of the neuroscientific data. Neuroscience and phenomenological analysis are not at odds—in fact, while they study what is nominally the same object, they do so in radically different but not competing modes—neuroscience grounding itself inside the spectrum of known-unknown, phenomenology grounding itself inside that of clear-obscure. Neuroscience expands the domain of what is given to intelligibility, the number of pos-

10 See Husserl, *Introduction*, 401: “The phenomenological reduction amounts to being constantly conscious of this fact and not passing between with naturalistic trains of thought where theory of knowledge is under investigation.”

sible of objects of knowledge, while phenomenology clarifies the concepts at our disposal in understanding or interpreting said domain. This provides a decent model for the interaction between neuroscience and the humanities, on which I will expand in the following section of this paper.

Understanding: concluding unscientific postscript

In place of a formal conclusion, I would like to draw the reader’s attention to an old distinction that comes out of the hermeneutical tradition via Wilhelm Dilthey: the distinction between explanation and understanding. Dilthey saw explanation as the mode of inquiry proper to natural science, and understanding as the mode of inquiry proper to the human sciences, or what we would today call the humanities. Explanation proceeds by the clarification of causal links in order to come up with law-based or rule-governed accounts of how phenomena arise. Understanding, on the other hand, takes human historical life into account, and attempts to integrate both the historicity and possible meanings of phenomena into an account. Explanation is “vertical,” insofar as causal relations determine from the ground up. Understanding is “horizontal,” insofar as it takes in our meaning-making horizons as indicative. We can see how this distinction might be helpful in glossing what the *neuro-* turn in the humanities is doing when it applies neuroscientific insights to objects of the humanities. If we take a set of insights that are at home in the natural sciences (and thus engage in the mode of explanation) and apply them to cultural objects (a novel, a mind, etc.) without first parsing our engagement with the methodological problem of moving from explanation to understanding, then we are simply treating cultural objects as objects of natural-scientific inquiry.

The interaction between explanation and understanding is complex, and often a particular analysis will involve both, whether it be a nominally scientific or nominally humanistic analysis. For an example of how these modes of analysis can and should interact, I would like to investigate one particular instance where the insights of phenomenology can be made available in the interpretation of neuroscientific research in a way that is beneficial to both.

Both cognitive neuroscience and phenomenology take up empathy as an object of study; as such, it seems like a good candidate for showing how the two might or might not be able to collaborate. Ralph Adolphs, a

researcher in neuropsychology, argues in “Emotion, Social Cognition, and the Human Brain” that understanding the actions and emotions of others means, in some sense, simulating them. The brain constructs social cognition, at least in part, by simulating the emotional state of the person being observed: we know what others experience because our brains reflect their sentiments.¹¹ This “insight” is based on the functioning of mirror neurons and the fact that certain brain areas are involved both when we experience certain stimuli and when we witness others undergoing the same stimuli. The automatic and covert operations of these brain areas in many instances, however, should give us pause in immediately adopting this explanation in the case of something like empathy in particular, or of social cognition in general. Contrary to the position both Adolphs and Antonio Damasio take, what is going on when mirror-neurons activate might not be that we are undergoing the experience of the other, albeit in a more superficial manner.¹² If we adopt a phenomenological standpoint from which to clarify and interpret the data at hand, we will come to a slightly different, yet immeasurably more fruitful interpretation. Adolphs and Damasio imply that, because the same neural networks are involved in both the first-personal experience of an emotion or of pain and in the empathetic understanding of that pain, at some level the experiences are the same, though less “intense.”

Phenomenologists have insisted since the beginning of the 20th century that empathy involves an analogization of one’s experience, and they have provided experiential and transcendental rationales for distinguishing between empathy, sympathy, and first-hand experience. These might seem like basic, unscientific, even uninteresting insights at first glance, yet they outstrip neuroscientific accounts insofar as they provide an internal, non-reductive account of that which is to be explained. In its transcendental modality, phenomenology gives an account of the conditions under which we speak of empathy (or freedom, or pain, or love, or fear, or religious *ekstasis*), as well as the conditions under which we might be given to speak of “imitation” or of the “experience of another.” In the kind of neuroscientific research mentioned above, a form of tautol-

11 Ralph Adolphs, “Emotion, Social Cognition, and the Human Brain,” in *Essays in Social Neuroscience*, eds. John T. Cacioppo and Gary G Berntson (Cambridge: MIT Press, 2004), 125.

12 Antonio Damasio, *Looking For Spinoza: Joy, Sorrow, and the Feeling Brain* (London: William Heinemann Press: 2003), 115.

ogy or question-begging arises from the experimental assumptions and explanatory framework—namely, this kind of research takes for granted that these things are constitutively and entirely in the brain, and so, even when dealing with social phenomena such as empathy, these explanations will always find a way to reduce the phenomenon to a brain structure, a procedure which may well be valid in some sense, but seems incomplete.

There is no reason to deny that the imitative brain structures at play in empathy constitute an innate human structure, but there is also no reason to elaborate from there that this structure is empathy, nor that “imitation” is *the* key to it, nor that because the structure is innate, there is no socio-cultural element to it. Not only has some recent neuroscientific research reified, individualized, and biologized what phenomenology would show to be a social fact,¹³ but the research on empathy as a whole has elided a few key phenomenological distinctions that seem to hold true. Firstly, the experience of empathizing is *different from* the experience of imitating another’s experience, regardless of the activation of mirror neurons. The qualitative experience of empathizing is distinguishable from a first-person experience of the same phenomenon, as it is distinguishable from imitating said experience imaginatively and from remembering said experience happening to oneself or another (as long as the memory does not reignite an emotion, which it might). Neuroscience is refined enough to distinguish between these instances, but Adolphs and Damasio make mirror neurons a sufficient condition for all of them. According to them, these things are more or less neurologically identical, despite how different they *feel*. That may be the case, but this just shows that there is more to these phenomena than their neurophysiological substrates. The final, most probative instance where phenomenology becomes instructive, if not indispensable, is the following question: If mirror neurons are active in empathetic experiences, if the same brain areas light up when I witness another in pain as when I myself am in pain, why does this mean that I experience it “in the second person,” as it were? It is absurd to reduce social, interpersonal, intersubjective structures, involv-

13 See the following study on implicit racial bias, which explains its results partly in terms of evolutionary biology, and without any reference to acculturation, which seems to be a rather important lacuna: Forgiarini *et al.* “Racism and the Empathy for Pain on Our Skin,” *Frontiers in Psychology* vol. 2 (2011), accessed on November 2, 2014. <http://dx.doi.org/10.3389/fpsyg.2011.00108>

ing multiple people and a social situation and a history and cultural ideas, to structures of individual neurophysiology. According to at least one study, the analysis of mirror neurons should shed light on intersubjectivity;¹⁴ instead, the experimental assumptions involved erode the very possibility of investigating intersubjectivity. This is why phenomenology, and the humanities in general, are of import to neuroscientific explanations of phenomena.

Neuroscience says too much and reduces it to too little. Although I have mostly concentrated on phenomenology, my statements can with little effort be generalised to the rest of the humanities, insofar as their insights do not attempt to contradict the facts while providing understanding. Interpreting the results of neuroscientific research as it pertains to higher-order phenomena requires a humanistic understanding of that which is to be explained, a refined conceptual toolkit for speaking about experience, culture, and interpersonal interaction that takes a different approach from that of neuroscience. I do not advocate such a requirement in order to discredit neuroscientific research, but in order to supplement it, steer the research in a productive manner, and perhaps refine the experiments that neuroscientists engage in. In the final analysis, neuroscience does not supplant humanistic scholarship, but enlarges the field of givenness that the humanities can engage with.



“AN EXPRESSLY NON-REDUCTIVE ACCOUNT OF THE SUBJECT INFORMED BY NEUROSCIENTIFIC INSIGHTS”

A brief response by Thomas Wormald

Mr. Samson and I fundamentally agree that a confrontation between the neurosciences and the humanities is a necessary and productive endeavour. Moreover, perhaps even an inevitable one and thus of equal importance—as I take the essential point of Samson’s argument to be—

14 Matthew Ratcliffe, “Phenomenology, Neuroscience, and Intersubjectivity,” in *A Companion to Phenomenology and Existentialism*, eds. Hubert Dreyfus and Mark A. Wrathall (Malden, MA: Blackwell, 2006), 329-45.

is the particular form this engagement takes. Of equal importance is the particular manner and care through which this endeavour is conducted, in regards to our mindfulness of the epistemological and ontological commitments, assumptions and limitations which must be recognized and negotiated when crossing and engaging different disciplinary registers. In this manner, the set of problems and concerns Samson foregrounds in his paper are of immense importance in their exhortation to think and reflect carefully on how we conduct and produce knowledge in an interdisciplinary field and the responsibilities inherent in negotiating these different disciplines.

Where Samson and I principally diverge is our interpretative differences concerning Catherine Malabou's work in *What Should We Do With Our Brain?*, and our respective evaluations of the success of her work as an exemplar of engagement between the humanities and the neurosciences. As evinced by one of the dominant leitmotifs operative in Samson's paper—that can be seen in notions such as first philosophy, hyper-empiricism and explanatory monism—what I take to be at stake in Samson's estimation, in regards to neuroscience and the naturalist epistemologies that subtend other neighboring disciplines, is the seductive reductionism of scientific explanation. That is, the allure that accompanies the notion that one can identify one substance or one cause or one set of predictable, naturalized laws, which all phenomena can be reduced to or explained by (or, as Samson writes, explained *away*). Essentially, the problem Samson identifies is the category error that occurs when such explanatory paradigms are applied to human beings. That is, when one exports the framework of natural science to explain human beings, one distorts ones object of inquiry in this case as much as one actually understands. Such is the fear of humanities writ large: science cannot account for the irreducibility of human experience.

Adducing Malabou here is, for me, quite strange. What Malabou is attempting to do is to articulate an expressly non-reductive account of the subject informed by neuroscientific insights. Specifically, Malabou's novel claim is that the sciences provide us with the means to do so: that science is no longer an enemy, but an ally in defending a robust account of human subjectivity. That is, while Malabou may say things like “you are your synapses,” what she is essentially doing is trying to make us see that our conventional understanding of “synapses,” “brain,” “program,” any of these traditionally understood scientific or fixed, causally mechanistic notions,

and the fear of determinism they evoke, needs to be substantially revised as borne out by the insights of contemporary science. While Malabou says ‘yes’ we are reducible to something like neuronal functioning, our brain is no longer a *reducible* or *reductive* entity in itself—as fundamentally plastic, it is an autopoietic economy of passivity and activity; we are not determined by program or design, but are both passively shaped by, and actively self-shaping through, our own singular histories and our interaction with our environs and one another. We are thus, in a paradox that undermines any determinism, reducible to an irreducibility. Malabou refuses to cede to either side of absolute naturalization or anti-naturalism, recognizing that the truth dialectically inheres in both. What Malabou wants to break down is essentially the “cold war” of science and philosophy that she sees as prohibitive to a full understanding of the material self. Malabou is interested in and affirms the complex, dialectical genesis of a processual subject, one in *becoming*, that is born of the struggle—and not the reductive continuity—between the biological and cultural, the neuronal and mental. This is the explanatory gap that Samson mentions, which Malabou takes to be the space of negation that opens up actual *material* freedom in the sense of plasticity. The critique of reductive, hyper-empirical, explanatory monism outlined here is germane and important, but it is misplaced in the instance of Malabou.

Samson is also critical of whether Malabou realizes her project in *What Should We Do With Our Brain?*, one which he defines as an “awakening of the consciousness of the brain”—a characterization with which I agree, but an assessment with which I do not. Samson writes:

for all its lip-service to neuroscience and its invocations of the terminology and relevant texts, [the actual text] has little to do with the initially announced project [being an awakening of the historicity of the brain] and instead has recourse to (deconstructive) textual hermeneutics.

Samson takes issue with a mixing of critical registers: an implied naturalistic focus that is effected or conducted through vaguely deconstructive strategies. This is coded negatively and I want to question why. This accusation is not entirely fair and can be allayed by keeping in mind how Malabou frames the text: through a gesture to Marx’s famous *Eigh-*

teenth Brumaire sentiment that humans make their own history, but they do not know that they do it and they do not do it exactly as they please. Translated into Malabou's terms, humans make their own brains, but they do not know that they do it and do not do it exactly as they please.

A productive way of situating Malabou's intervention in or engagement with neuroscience is by keeping in mind its self-avowed parallels with Marx. That is, just as modes of production are ideologically naturalized in Marx's analysis—they are necessary, determined, inevitable, unable to be changed—Malabou argues that a similar naturalization occurs with popular understandings of the brain that engender equally pernicious effects, creating a naïve consciousness or attitude that construes oneself, and possibly one's world, as ultimately determined and unchangeable; without a genuine future or possibility. I think that this is a perfect example of how the humanities can use its own conceptual tool-kit and resources to complicate, check, or bear upon discourses of science: it can bring a sense of critical reflexivity and sensitivity to how the ostensibly 'neutral' discoveries and discourses of "objective" science play out and can be mobilized in the world to status quo serving political, social and cultural ends. In this way, I take Malabou's project to be precisely what Samson characterizes as a productive way for a philosopher to approach the findings of neuroscientific research: as a field of givenness that expands, troubles, and pushes us to re-conceptualize what it means to be human—particularly as it affords robust resources to defend and bolster the *material* reality of human possibility, a future and transformative political action.