

Closing the Gap? Some Questions for Neurophenomenology¹

Tim Bayne
University of Oxford and St. Catherine's College
Manor Road
Oxford OX1 3UJ
United Kingdom
tim.bayne@gmail.com

This paper was published in (2004) *Phenomenology and the Cognitive Sciences*, 3/4: 349-64.
Please consult the published version for purposes of quotation.

In his 1996 paper “Neurophenomenology: A methodological remedy for the hard problem”, Francisco Varela called for a union of Husserlian phenomenology and cognitive science. Varela’s call has not gone unanswered, and recent years have seen the development of a small but growing literature intent on exploring the interface between phenomenology and cognitive science. Indeed, this very journal is, in large part, a testimony to the rise of neurophenomenology. But despite these developments, there is still some obscurity about what exactly neurophenomenology is.² What are neurophenomenologists trying to do, and how are they trying to do it? To what extent is neurophenomenology a *distinctive* research programme? My aim in this paper is not so much to answer these questions as to provoke neurophenomenologists into answering them.

¹ I would like to thank Dave Chalmers, Barry Dainton, David Mackenzie, Evan Thompson, Shaun Gallagher and Antoine Lutz for their many helpful comments on this paper. This paper began life as a review of *Naturalizing Phenomenology*, and I am grateful to the editors of this journal for permitting me to transform it into its present form.

² For another perspective on neurophenomenology see Gallagher (1997) and Gallagher and Varela (2001).

Getting situated

What is the goal of neurophenomenology? According to Varela, neurophenomenology seeks to address the hard problem.

Neuro-phenomenology is the name I am using here to designate a quest to marry modern cognitive science and a *disciplined approach* to human experience, thus placing myself in the lineage of the continental tradition of phenomenology. My claim is that the so-called hard problem ... can only be addressed productively by gathering a research community armed with new pragmatic tools enabling them to develop a science of consciousness (Varela 1996: 330; emphasis in original).

Later in the article Varela says that neurophenomenology is intent on closing the explanatory gap. The explanatory gap is the gap between the phenomenal – or “what it’s like” – character of experience and the physical nature of the brain/body (Levine 1983). It is the explanatory gap that makes the hard problem hard. If one could close the explanatory gap one would have solved the hard problem. Neurophenomenology seeks to close the gap by showing that phenomenality can be “integrated into an explanatory framework where every acceptable property is made continuous with the properties accepted by the natural sciences” (Roy et al 1999: 1f).³

How do neurophenomenologists propose to close the explanatory gap? In answering this question neurophenomenologists often refer to a distinctive method for the first-person study of consciousness. This is where the phenomenology comes in. By drawing on the tools developed by Husserl and other phenomenologists, it is hoped that first-person methods in the study of consciousness can become as rigorous as third-person methods. We need better descriptions (models) of first-person data, and, so neurophenomenologists claim, (only) the methods of Husserl and his followers can provide this.

Is this *all* we need to do in order to close the gap? Can we close the gap *merely* by developing better models of phenomenality and better models of neuronal activity? Some of those sympathetic to neurophenomenology seem to suggest that it might be. According to Borrett, Kelly and Kwan,

...the right relationship between phenomenology and brain science is that of data to model: brain science is ultimately concerned with explaining the way the physical processes of the brain conspire to produce the phenomena of human experience; insofar as phenomenology devotes itself to the accurate description of these

³ One might wonder what it is for a property to be continuous with the properties accepted by the natural sciences.

phenomena, it provides the most complete and accurate presentation of the data that ultimately must be accounted for by models of brain function (2000: 213; see also Marbach 1999)

On the whole, however, neurophenomenologists suggest that neurophenomenology involves more than just neuroscience *and* phenomenology. Exactly what that something more is, I think, an open question. Varela speaks of changing the “entire framework within which the [hard problem] is discussed” (Varela 1996: 340). The idea seems to be that neurophenomenology solves the hard problem by undermining the assumptions that generate it.

Instead of finding ‘extra ingredients’ to account for how consciousness emerges from matter and brain, my proposal reframes the question to that of finding meaningful bridges between two “irreducible” phenomenal [sic: I presume that Varela is referring to experience and the brain] domains. In this sense neurophenomenology is a potential solution to the hard problem by casting in [sic] an entirely different light on what ‘hard’ means. (1996: 340)

Exactly how Varela hopes to cast an entirely different light on “what ‘hard’ means” is unclear to me, but the idea seems to have something to do with the notion of a meaningful bridge between the first-person science of phenomenology and the third-person sciences of the brain. The second half of this paper will explore this idea in detail.

So, at a general level one can discern two strategies at work in neurophenomenology. The first strategy employs the techniques of (Husserlian) phenomenology to arrive at better first-person models of phenomenality. Call this the *descriptive strategy*. The second strategy attempts to bridge phenomenological models and neuroscientific (or, more broadly, cognitive scientific) models. Call this the *bridging strategy*. Although neurophenomenologists have not clearly distinguished the descriptive and bridging strategies, it seems to me that they are independent. One could hold that (Husserlian) phenomenology is uniquely positioned to describe experience while being agnostic about whether it has anything original to say about how to bridge the gap between such descriptions and neuroscience. Alternatively, one could hold that neurophenomenology can close the explanatory gap while being agnostic about whether it is in privileged position to describe phenomenality. I will examine the descriptive strategy and the bridging strategy in that order.

The descriptive strategy

At the heart of neurophenomenology is the claim that Husserlian phenomenology has a unique and privileged method of describing the first-person nature of consciousness.

Phenomenology is *unique* in that it is importantly different from the standard first-person methodologies employed in consciousness studies; and it is *privileged* in that it is more rigorous than such methodologies.

I have my doubts about both claims. I say this with some tentativeness, for my knowledge of the phenomenological tradition is limited.⁴ But let me explain why I am inclined to think that the phenomenological method is neither unique nor privileged.

It is often said that the phenomenological reduction is not to be confused with introspection. Varela states that phenomenology “goes beyond mere introspection” (Varela 1996: 338); Thompson, Noë and Pessoa say that the phenomenological method of reflection is “fundamentally different in its procedures and aims from introspection” (1999: 571); and Marbach claims that the neglect of Husserl in cognitive science is linked to the “widespread misunderstanding, or misrepresentation, of his method of *reflective* phenomenology of conscious awareness as being a study of conscious awareness *through introspection*” (1993: 15; emphasis in original).

How does phenomenological reflection differ from introspection? I have yet to find a clear answer to this question. The editors of *Naturalizing Phenomenology* describe phenomenology (in the narrow, Husserlian sense) as “an enterprise primarily dedicated to the careful establishment, through first-person description and analysis, of phenomenological data understood as what we are really aware of, as opposed to what we believe we are aware of” (Roy et al 1999: 18f). This might be true, but it fails to distinguish phenomenological reflection from introspection, for the goal of introspection is also to describe what we are really aware of (as opposed to what we believe we are aware of).

Varela’s explication of the difference between phenomenological reduction (PhR) and introspection is no more enlightening.

Phenomenology does share with introspectionism an interest in the reflective doubling as a key move of its approach to phenomena. But the two attitudes part company. In PhR the skill to be mobilised is called bracketing for good reasons, since it seeks precisely the opposite effect of an uncritical introspection: it cuts short our quick and fast elaborations and beliefs, in particular locating and putting in abeyance what we think we ‘should’ find, or some ‘expected’ description. (1996: 338f.)

⁴ For discussion of the phenomenological method see Varela (1996), and Depraz, Varela and Vermersch (2000).

Surely both the phenomenological reduction and introspection can be carried out either critically or uncritically, carefully or carelessly. I find nothing here that reveals how the phenomenological reduction differs from introspection as such.

One point of contrast between the introspectionist and the phenomenologist is that the latter is, I take it, interested in essences. I am not sure what a Husserlian essence is, but I suspect that it has something to do with necessity; the analysis of essences, I presume, involves necessary truths. And here there does seem to be an important difference between phenomenology and introspection, for introspection does not, I think, give one access to necessary truths (as necessary truths). One can justify claims about the nature of one's own consciousness on the basis of introspective, but it is at best unclear how introspection can justify claims about the necessary structure of consciousness in general. Indeed, it is not clear how introspection could justify claims about the necessary structure of *one's own* consciousness. Introspection may be able to tell me that my consciousness happens to be unified, but can it tell me that my consciousness must be unified? If so, it is not obvious how. So, in so far as phenomenologists are attempting to uncover the essential structure of experience it is implausible to suppose that they could rely on (mere) introspection alone.

The foregoing suggests that perhaps phenomenological reflection involves the use of intuitions of the kind that are employed in conceptual analysis and thought-experiments. Here the object of analysis is not an experience but a concept (albeit, perhaps, the concept of experience). There is some support for this view among Husserl's commentators. D'Amico, for instance, claims that Husserlian intentional analysis is "pure conceptual analysis" (D'Amico 1999: 15).⁵

Unfortunately, this view does not seem to be shared by neurophenomenologists. Thompson, Noë and Pessoa are particularly keen to distinguish the phenomenological method from intuitions as they are used in thought-experiments.

....phenomenological claims about essences and "eidetic" necessities are not grounded on (typically underdescribed) "intuition pumps" of current philosophy of mind. On the contrary, ideation through imaginative variation requires sustained

⁵ Crowell comments: "It appears that only ignorance informs the view that phenomenology's results are nothing but conceptual analysis. One might more justly say that there is conceptual analysis only because there is phenomenology, even though its practitioners don't recognize themselves as phenomenologists" (2002: 441). Either way, Crowell seems to agree that there isn't much to differentiate phenomenology from conceptual analysis properly conducted.

attention to and careful description of the phenomena in all their ramifications. Furthermore, the phenomena must be attended to not simply as particularities, but as fields of possibilities defined by certain forms or structural invariants. Finally, success in this endeavour depends on unprejudiced reflection, that is, on suspending or refraining from making use of any judgments about how things might be apart from our experiences of them. (1999: 573, n. 11)

Perhaps we should not place too much weight on a point that Thompson, Noë and Pessoa consign to a footnote, but from what they say I am not at all clear how “ideation through imaginative variation” differs from the appeal to thought-experiments and intuition pumps. What exactly is it to attend to a phenomenon as a “field of possibilities defined by certain forms or structural invariants”? I thought the goal was to arrive at (justified) claims about the structural invariants of consciousness: how could attending to something as defined by a structural invariant enable one to justify a claims about what is or is not a structural invariant of consciousness? As to the claim that unprejudiced reflection is necessary, this may be true, but it does not advance the project of understanding how ideation through imaginative variation differs from the use of intuition pumps.

If neurophenomenologists are interested in essences there is little evidence of it in what they actually do. Take, for instance, the work by Varela (1999) and van Gelder (1999) on the specious present. I take these authors to be asking the following question: what explains the structure of phenomenal temporality that humans tend to enjoy in normal states of consciousness? As far as I can tell, there is nothing in this work that gets at the *essential* structure of consciousness as such, or even at the essential structure of human consciousness.

More generally, I can discern little evidence of any of the “technical developments of Husserlian phenomenology” (Roy et al 1999: 21) at work in neurophenomenology. It seems to me that the methods for collecting first-person data employed by neurophenomenologists are much the same as those employed elsewhere in the study of consciousness.

Consider, for instance, Thompson, Noë and Pessoa’s criticisms of Dennett on the phenomenology of parafoveal vision. The question with which Thompson, Noë and Pessoa are concerned is this: What is the content of visual phenomenology when looking at wallpaper whose pattern is a regular array of hundreds of identical images of Marilyn Monroe? Dennett had suggested that it would seem as though one were seeing *detailed* images of hundreds of Marylins (Dennett 1991: 354f.). Thompson, Noë and Pessoa disagree: “Although you do seem to see all the detail in the sense that the wall seems to

you to be covered with hundreds of identical Marilyns, you do not seem to see each Marilyn equally well. At any given moment of your perception, the Marilyns straight ahead seem clear as day, while those off to the side appear less distinct, and those in the periphery seem barely noticeable" (1999: 187). Dead right, but as far as I can see this claim is justified by an appeal to – mere, albeit careful – introspection.

Consider also the recent study by Lutz et al, billed by Lutz as an explicit implementation of the neurophenomenological program (Lutz 2002; Lutz et al. 2002). The study involved the perception of a 3-D object arising from an autostereogram.

The task began when the subjects fixed [sic: fixated] a dot-pattern containing no depth information. After an auditory signal, the subjects were asked to fuse two little squares at the bottom of the screen and to remain in this eye position for seven seconds. At the end of this preparation period, the random-dot pattern was changed to a slightly different random-dot pattern with binocular disparity (autostereogram). Subjects were readily able to see a 3D illusory geometric shape (depth illusion). They were instructed to press a button with their right hand as soon as the shape had completely emerged. This ended the trial, after which subjects gave a brief verbal report of their experience. In their reports the subjects used phenomenal invariants (or categories) found and stabilized during the training session. (Lutz 2002: 140)

What did these training sessions involve? How was the phenomenological reduction employed here? Lutz informs us that the gesture of reduction was either self-induced by subjects familiar with it, or induced through open questions such as "what did you feel before and after the image appeared?" I must confess that I do not find much evidence here of "technical developments" of any kind. I can see only two differences between this experimental protocol and that of standard experimental work on consciousness. First, these subjects had more exposure to the stimulus than they normally have in this kind of experimental work; second, they were required to report their phenomenal states by using phenomenal categories that they themselves had developed. While both departures from normal experimental protocol are interesting, I don't see any evidence here of a great methodological leap forward. In particular, I don't see any evidence that Lutz's subjects were no longer employing mere introspection in reporting their experiences.

As we have seen, when writing on the topic neurophenomenologists insist that there is a crucial difference between phenomenology (rigorous and scientific) and introspection (sloppy and unscientific). Yet, in other contexts, they seem remarkably relaxed about the distinction. Consider, for example, the structure of *The View from Within*, a recent collection of essays examining first-person methods in consciousness studies edited by

Varela and Shear (1999). This volume is divided into a number of parts, the first of which is entitled 'introspection' and the second is entitled 'phenomenology'. Yet one of the papers included in the phenomenology section is Andrew Bailey's "Beyond the Fringe: William James on the transitional parts of the stream of consciousness." Was James, I wondered, a (crypto) phenomenologist? Did he arrive at his account of the structure of the stream of consciousness on the basis of 'mere introspection', or was he – as the inclusion of Bailey's paper in the Phenomenology section of the volume might suggest – a secret proponent of the Husserlian method? As far as I can tell, James was nothing but a good old-fashioned introspectionist. Indeed, Bailey informs us that James was no fan of Husserl's, and even delayed the publication of his *Logische Untersuchungen* in the United States.

However exactly the phenomenological reduction differs from introspection, neurophenomenologists claim that it's more reliable. With the reduction in hand, Varela tells us, we can leave behind the squabbles that bedevilled the use of introspection in pre-phenomenology days (Varela 1996: 338). Again, I have my doubts. Consider just one case, that of time consciousness. Husserl devoted much effort to the analysis of phenomenal temporality, yet he changed his views on the topic in fundamental ways.⁶ Furthermore, there seem to be remarkable parallels between C. D. Broad's analyses of time-consciousness – who, I presume, was using mere introspection – and Husserl's (see Dainton 2000). It looks very much as though there may not be anything particularly privileged about the phenomenological reduction.⁷

My central point is that more needs to be said about how the phenomenological method differs from other first-person methods in psychology and philosophy. In what ways does phenomenological reduction depart from (mere) introspection? How does "ideation through imaginative variation" differ from the appeal to thought-experiments? And what

⁶ Citing Blattner (1999), Brough (1989) and Miller (1984), van Gelder says that there is a remarkable level of agreement about what Husserl's account of time-consciousness was. But in fact there are important points of disagreement between Brough and Miller on Husserl's account of time-consciousness. I am indebted to Shaun Gallagher (personal communication) here. See also Gallagher 1998: 60f.

⁷ The editors of *Naturalizing Phenomenology* remark on the fact that although "Jackendoff makes no reference to Husserlian phenomenology, it is quite surprising to see how closely his analysis parallels those [sic] of Husserl." (Roy et al. 1999: 21.) Doesn't this suggest that there is nothing special about Husserlian phenomenology?

reason do we have for thinking that the results of phenomenological reduction are more reliable than those of 'mere introspection'?

The bridging strategy

I turn now to the second of the two neurophenomenological strategies that I earlier identified, that of building a meaningful bridge between phenomenology and neuroscience (or, more broadly, cognitive science). There are two questions here for the neurophenomenologist: *What* would it take to have a meaningful bridge between phenomenology and neuroscience? *How* does one go about building such a bridge? Let me begin with the second question. Although this might seem a counter-intuitive order in which to take these questions, we can use what neurophenomenologists say about how to build a bridge as a guide to what they think it takes to be a bridge.

There are a number of neurophenomenological accounts of how to bridge the gap between neuroscience and phenomenology. Although I am not sure how these accounts are meant to be related, it seems fairly clear that they are all connected to what Varela called the working hypothesis of neurophenomenology: the *methodology of reciprocal constraints* (Varela 1996: 343; see also van Gelder 1999: 246). What exactly does the *methodology of reciprocal constraints* (MRC) involve?

MRC as reflective equilibrium

At first glance MRC sounds like the well-known notion of reflective equilibrium; Thompson, Noë and Pessoa suggest that the two notions are interchangeable (1999: 195), while Lutz claims that "the key point" behind MRC is that phenomenological evidence and cognitive evidence are granted an "equal importance and therefore need the same attention" (Lutz 2002: 135). Varela, however, explicitly distinguished MRC from reflective equilibrium: the former, he says, is more precise and more demanding than the latter (Varela 1996: 344). Unfortunately, he does not expand on this claim.

In philosophy the phrase "reflective equilibrium" is used to refer to the idea that intuitions about particular cases should be brought into alignment – reflective equilibrium – with intuitions about general principles. For example, in developing an account of morality one ought to give epistemic priority neither to one's particular moral intuitions nor to one's intuitions about moral principles but instead accord both kinds of intuitions equal weight. In philosophy the entities that are to be brought into equilibrium are intuitions, but the notion of reflective equilibrium can be extended to apply to all forms of data.

There are two important things to note about reflective equilibrium. First, it is an *epistemic* principle not an *explanatory* principle. To claim that two sets of data ought to be brought into equilibrium is to make no claims about the explanatory relations between them.

Consider, for example, the case of testimony. I might think that Amy's claims and Aileen's claims ought to be subject to reflective equilibrium without thinking that either sort of claim explains the other. Reflective equilibrium, as such, bridges no explanatory gaps. To the extent that MRC involves positing reciprocal *explanatory* relations between phenomenology and neuroscience it must go beyond reflective equilibrium.

The second point to note is that in order for reflective equilibrium between two data sets (A and B) to have any bite it must be possible for A-statements and B-statements to be inconsistent. It is clear that intuitions about particular moral judgments and general moral judgments can be inconsistent; after all, they draw on a common pool of concepts. The intuition that is permissible to lie in such-and-such a context is inconsistent with the intuition that lying is always wrong. But how can phenomenological statements be inconsistent with neuroscientific statements? The two kinds of statements draw on different pools of concepts. A priori, there are no entailment relations between descriptions of phenomenal states and descriptions of neuroscientific states.

In order for reflective equilibrium between phenomenology and neuroscientific to get off the ground we need *bridging principles* that link phenomenal data and neuroscientific data. Given that these principles cannot be established a priori they must be established a posteriori on the basis of correlations between phenomenal data and neuroscientific data. Exactly how such bridging principles can be discovered is, I think, a topic that deserves more attention than it has received to date.⁸

⁸ The fact that bridging principles can only be established a posteriori is related to the vehicle-content distinction (see Dennett 1991; Gallagher 1997; Millikan 1993; Hurley 1998). Neuroscience studies phenomenal states qua vehicles, phenomenology studies them qua content-bearing states. (Or, as we might also put it, neuroscience studies experiences qua syntactic entities, phenomenology studies them qua semantic entities.) Arguably, the relation between vehicle and content is contingent, and thus can only be discovered a posteriori. Studying experiences at the level of vehicles won't tell you what their phenomenology is. This is fairly obvious when it comes to such phenomenal states such as color experiences. No one thinks that the neural state that is (or supports, or causes, or grounds, or whatever) a red experience must itself be red. And no one thinks that the neural state that is (or supports, or causes, or grounds, etc) a visual experience of a square barn must itself be square. And if it is square, it isn't an experience of a square barn because it itself is square. To put the point as it is often put, the properties of the contents of

Another worry about unpacking MRC in terms of reflective equilibrium concerns the relationship between reflective equilibrium and conceptual autonomy. Can one consistently hold that two domains ought to be brought into reflective equilibrium with each other and also hold that each domain is conceptually autonomous? I'm not sure that one can. Doesn't the rejection of the claim that two domains are conceptually autonomous simply amount to the thought that they ought to be brought into reflective equilibrium with each other? Yet neurophenomenologists seem intent on endorsing the autonomy of phenomenology while holding that its deliverances ought to be brought into equilibrium with those of neuroscience. According to Thompson, Noë and Pessoa, upholding the conceptual autonomy of the personal level means "treating our understanding of ourselves as conscious perceptual subjects as a distinctive form of understanding, one that can be brought into 'mutual accommodation'; ... or 'reflective equilibrium' ... with cognitive science" (1999: 195). At least on the face of things, there seems to be a tension between the commitment to autonomy and the commitment to equilibrium. What does it mean to describe the personal level as autonomous if autonomy is compatible with reflective equilibrium?

MRC as a heuristic strategy

A second (and closely related) conception of MRC conceives of the methodology in heuristic terms, according to which phenomenological data can be used as a guide in the discovery of neuroscientific data, and vice-versa. This reading of MRC derives primarily from the work of Lutz et al. (2002), discussed above. As mentioned, the subjects in Lutz's study reported their experiences in terms of three phenomenological categories (or "clusters"): steady readiness, fragmented readiness, and unreadiness. Lutz and collaborators used these categories to discover previously undetected structure in the neuronal states of their subjects. The important point here is that the subject's *phenomenal* reports guided the analysis of the *neuroscientific* data.

This is, I think, a fascinating result. Phenomenal data are being used as a heuristic (or guide) in the investigation of the neuroscientific structure of the mind. And while Lutz and co-workers didn't use neuronal data as a heuristic in the discovery of phenomenal structure, I can see nothing to stop a research team attempting to employ such a method. I think this is an important study, but does it shed any light on the explanatory gap? Lutz and his co-authors think so: "This simple case study ... already illustrates how fertile this

consciousness (that's the phenomenal part) need not be identical to the properties of the vehicles of consciousness (that's the neuroscientific part).

approach could be to identify biophysical properties and to understand their relation to experience” (Lutz 2002: 17). The first claim is justified but the second claim is not. This study no more illuminates the relationship between biophysical properties and phenomenal properties than any other work on the neural correlates of consciousness does; which is to say that it doesn’t. The gap between the events, processes and structures discovered by neuroscience and the events, processes and structures open to phenomenology remains as wide as ever.

MRC as reciprocal causation

A third possible reading of MRC is *causal*. In their paper “Radical embodiment: neural dynamics and consciousness” Varela and Thompson (2001) argue that there are two kinds of causation in the brain: ‘upward’, or local-to-global causation, and ‘downward’, or global-to-local causation. As I read them, Thompson and Varela hold that this kind of reciprocal neural causation is also a form of phenomenal–neural causation. There are, they say, “reciprocal causal-explanatory relationships between neural events and conscious events when the latter are conceived of as order parameters of large-scale brain dynamics” (Thompson and Varela 2001: 421).

There is something to be said for thinking of MRC in causal terms. For one thing, it appears to offer some hope of closing the explanatory gap in so far as causal relations are explanatory. On the other hand, MRC is meant to be a methodological thesis – that’s what the ‘M’ stands for – and this is slightly at odds with the causal reading of it. No matter; perhaps MRC is merely misdescribed. The central issue is whether neurophenomenologists are, or should be, interested in identifying reciprocal causal relations between phenomenal states and neuronal states.

Clearly the move from neural-neural causation to neural-phenomenal causation rests on a prior identification of neural states with phenomenal states. Although the above quotation from Varela and Thompson seems to endorse such an identification, it also runs counter to their very deeply anti-reductionist leanings. In numerous places Varela rejects any form of the identity theory, and Thompson has urged that to identify the contents of neural states with the contents of phenomenology is to commit something like a category mistake (Noë and Thompson, 2004). So, I am puzzled about how the Thompson-Varela argument for neural-phenomenal causation goes. It *looks* like it rests on some kind of (token) identity theory, but such a commitment seems to be inconsistent with the anti-reductionist tenor of neurophenomenology.

The identification of phenomenal states with neural states is also at odds with Varela and Thompson’s embedded (embodied, situated) approach to the mind. How could a

conscious event be an order parameter of large-scale brain dynamics if there is no such thing as a minimal internal correlate whose intrinsic properties are sufficient to produce conscious experience? According to proponents of the embedded approach to phenomenology, phenomenal content is had by an active (embodied, embedded, situated, etc.) animal. If that is right, then the vehicles of phenomenology are not identical to any part of an animal, even if that part is a global neural state. (Indeed, on the most radical embedded approaches to consciousness one should not even say that the animal as a whole is conscious, rather, it is the animal-in-its-environment that is the proper subject of consciousness.) In short, Varela and Thompson's case for reciprocal causal relations between phenomenal states and neuronal states seems to be inconsistent with their rejection of identity theories of consciousness and with their commitment to an embodied account of cognition.

Finally, we might note that merely establishing *that* there are causal relations does not suffice to close the explanatory gap. After all, Descartes – the archetypical non-naturalist – was more than willing to think that there is causal traffic between mind and matter. The challenge, it seems to me, is to make such traffic *intelligible*. More specifically, the challenge for the neurophenomenologist is to explain how reciprocal causal relations might hold within a naturalistic (but non-reductionist) account of the mind. As Lutz rather disarmingly puts it, “The introduction of reciprocal causation might be perceived as a subtle expression of dualism” (Lutz 2002: 150). Indeed it might.⁹

Isomorphism and Generative Passages

I turn now from the question of *how* to build an explanatory bridge between phenomenology and neuroscience to the question of *what it is* to build such a bridge. What would it take to close the explanatory gap?

The neurophenomenological answer to this question is, I think, bound up with the neurophenomenological account of what it is to naturalize a domain. That account gives an important role to mathematics. The editors of *Naturalizing Phenomenology* write: “the problem of naturalizing a Husserlian description amounts to articulating its mathematical reconstruction with relevant lower-level natural sciences, neurobiological disciplines being of course of primary importance” (Roy et al 1999: 63). The view, I take it, is that one

⁹ But perhaps neurophenomenology's commitment to naturalism is not unequivocal. Lutz writes of developing a model that will allow mental and natural properties to coexist without contradiction (Lutz 2002: 28) – thereby suggesting that mental properties are not natural properties.

can close the explanatory gap (only?) by developing mathematical models that involves variables that can refer to either phenomenal states (structures, events, etc) or neurophysiological states (structures, events, etc). This notion of a formal model that can be applied to both phenomenal and neurophysiological states has been dubbed a “generative passage” (see Varela 1997; Lutz 2002). A generative passage “constitutes an intermediate and neutral level in which both the experiential and biophysical levels can be expressed” (Lutz: 2002: 28).

I have three worries about employing the notion of a generative passage to close the explanatory gap. The first worry concerns the difference between the concept of a generative passage and that of a phenomenal-neuronal isomorphism. That the two notions are meant to be distinct is clear from the account of a generative passage provided by the editors of *Naturalizing Phenomenology*:

A more demanding approach [than the isomorphism strategy] will require that the isomorphism idea be taken one step forward to provide the passage where the mutual constraints not only share logical and epistemic accountability, but are further required to be operationally *generative*, that is to say, to be in a position to generate in a principled manner eidetic descriptions that can directly link to explicit processes of biological emergence. For this to happen at least both sides of the wavy line [the line dividing phenomenological descriptions and neurobiological descriptions] must be joined at a level of description sufficiently abstract that it rightly belongs to both sides at the same time (Roy et al 1999: 68).

It is completely unclear from this passage how the notion of a generative constraint differs from that of an isomorphism. I take it that an isomorphism between phenomenal states and neuronal states holds if and only if there is a formal model – that is, an abstract level of description – that can be applied to both domains.

Consider two important exercises in neurophenomenology - Varela’s (1999) and Van Gelder’s (1999) accounts of time consciousness. It seems to me that both accounts assume some form of isomorphism. This is, I think, particularly clear in Varela’s analysis, for he claims that the “various [neural] components require a *frame or window of simultaneity that corresponds to the duration of the lived present*” (1999: 272; original italics). What is a correspondence if not an isomorphism? I also detect isomorphism at work in Varela’s use of diagrams, which suggest that the cognitive present has the same phase structure as neuronal synchrony (see especially Varela 1999: 276).

My second worry with the notion of generative passage concerns the claim that we might be able to apply a single formal system to both neuroscientific and phenomenal states. If

I've understood the strategy correctly, the goal is to apply dynamical systems models to both phenomenology and neuroscience. But is it really likely that we will be able to develop dynamical models of first-person data? Varela provides some graphs of Husserl's analysis of time-consciousness, but they hardly amount to a formal model. I would not dismiss the possibility of a dynamical model of first-person data, but I do find it hard to imagine that we will ever be in a position to construct a decent mathematical model of the stream of consciousness. But perhaps I am being unduly pessimistic here.

My third worry is perhaps the most serious. Would a formal model that could be applied to both phenomenal events and neural events close the explanatory gap? I think not. We might be able to *predict* phenomenal states on the basis of neural data, but the explanatory itch would remain. We would still need to explain why some neural states have a phenomenology, and why particular neural states have the particular phenomenology that they do. Even with a complete formal model of both the biological mind and the phenomenal mind, it seems to me that the relations between the two would remain as mysterious as ever.

The reason for this, I think, is that formal models can only capture the *structure* of a domain; they cannot capture its intrinsic nature. Those who think that the hard problem is hard do so because they think that phenomenal character – the “what it's like” of experience – cannot be fully captured by structural descriptions. (Perhaps such folk are *wrong* to hold this view, but that is a separate issue.) To the extent that one is at all gripped by the explanatory gap, one will be inclined to doubt that it could be bridged by mathematical models. There is certain irony in the fact that neurophenomenologists seem to be assuming a functionalist account of phenomenology very much like the ones they frequently disparage.

Conclusion

Varela's description of neurophenomenology as a radically new research tradition for consciousness studies seems to me to be somewhat over-stated. As far as I can tell, neurophenomenologists have no recipe for closing the explanatory gap. In making this claim I do not in any way mean to dismiss the work done under the neurophenomenology label. It seems to me that Lutz's use of first-person data to guide the analysis of brain dynamics, for example, is an excellent model of how the science of consciousness ought to proceed. But neither it nor any other experimental work conducted by neurophenomenologists lives up to the bold claims made by its proponents. How it is that anything so remarkable as a state of consciousness comes about as a result

of irritating nervous tissue is still as remarkable as the appearance of Djin when Aladdin rubbed his lamp.

References

- Blattner, W. 1999. *Heidegger's Temporal Idealism*. Cambridge: Cambridge University Press.
- Borrett, D., Kelly, S. & Kwan, H. 2000. Phenomenology, dynamical neural networks and brain function. *Philosophical Psychology* 13/2: 213-228.
- Brough, J. 1989. Husserl's phenomenology of time-consciousness. In J.N. Mohanty and W.R. McKenna (eds). *Husserl's Phenomenology: A Textbook*. Lanham, MD: University Press of America.
- Chalmers, D. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2/3: 200-219.
- Crowell, S. 2002. Is there a phenomenological research program? *Synthese*, 131: 419-444.
- D'Amico, R. 1999. *Contemporary Continental Philosophy*. Boulder, CO: Westview Press.
- Dainton, B. 2000. *Stream of Consciousness: Unity and Continuity in Conscious Experience*. London: Routledge.
- Dennett, D. 1991. *Consciousness Explained*. Boston, MA: Little Brown and Company.
- Depraz, N., Varela, F. & Vermersch, P. 2000. The gesture of awareness. In M. Velmans (ed). *Investigating Phenomenal Consciousness* (pp. 121-136). Amsterdam: John Benjamins.
- Gallagher, S. 1997. Mutual enlightenment: Recent phenomenology in cognitive science. *Journal of Consciousness Studies* 4/3: 195-214.
- Gallagher, S. 1998. *The Inordinance of Time*. Evanston, IL: Northwestern University Press.
- Gallagher, S. and Varela, F. 2001. Redrawing the map and resetting the time: Phenomenology and the cognitive sciences. In S. Crowell, L. Embree and S. Julian (eds). *The Reach of Reflection: Issues for the Phenomenology's Second Century*. Center for Advanced Research in Phenomenology.
- Levine, J. 1983. Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64: 354-61.
- Lutz, A. 2002. Toward a Neurophenomenology of generative passages: A first empirical case study. *Phenomenology and the Cognitive Sciences* 1: 133-167.
- Lutz, A., Lachaux, J-P, Martinerie, J., & Varela, F. 2002. Guiding the study of brain dynamics by using first-person data: Synchrony patterns correlate with ongoing

conscious states during a simple visual task. *Proceedings of the National Academy of Sciences USA*: 99: 1586–1591.

Marbach, E. 1993. *Mental Representation and Consciousness: Towards a Phenomenological Theory of Representation and Reference*. Dordrecht: Kluwer.

Marbach, E. 1999. Building materials for the explanatory bridge. In F. Varela and J. Shear (eds.) *The View from Within* (pp. 252-257). Exeter: Imprint Academic.

Miller, I. 1984. *Husserl, Perception and Temporal Awareness*. Cambridge, MA: MIT Press.

Millikan, R. (1993) Content and vehicle. In N. Eilan, R. McCarthy, and B. Brewer (eds). *Spatial Representation*. Basil Blackwell: Oxford.

Noë, A. and Thompson, E. (2004) Are there neural correlates of consciousness? *Journal of Consciousness Studies* 11 (1): 3-28.

Roy, J-M, Petitot, J., Pachoud, B. & Varela, F. 1999. Beyond the gap: An introduction to naturalizing phenomenology. In J. Petitot, F. Varela, B. Pachoud & J-M Roy (eds). *Naturalizing Phenomenology* (pp.1-80). Stanford, CA: Stanford University Press.

Thompson, E., Noë, A. & Pessoa, L. 1999. Perceptual completion: A case study in phenomenology and cognitive science. In J. Petitot, F. Varela, B. Pachoud & J-M Roy (eds). *Naturalizing Phenomenology* (pp. 161-95). Stanford, CA: Stanford University Press.

Thompson, E. and Varela, F. 2001. Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences*, 5/10: 418-25.

van Gelder, T. 1999. Wooden iron? Husserlian phenomenology meets cognitive science. In J. Petitot, F. Varela, B. Pachoud & J-M Roy (eds). *Naturalizing Phenomenology* (pp. 245-65). Stanford, CA: Stanford University Press.

Varela, F. 1996. Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies* 3/4: 330-49.

Varela, F. 1997. The naturalization of phenomenology as the transcendence of nature: Searching for generative mutual constraints. *Alter: Revue de Phénoménologie*, 5: 355-385.

Varela, F. 1999. The specious present: A neurophenomenology of time consciousness. In J. Petitot, F. Varela, B. Pachoud & J-M Roy (eds). *Naturalizing Phenomenology* (pp. 266-312). Stanford, CA: Stanford University Press.

Varela, F. and Shear, J. (eds). 1999. *The View from Within*. Exeter: Imprint Academic.