

Alva Noë and Evan Thompson

Sorting Out the Neural Basis of Consciousness

Authors' Reply to Commentators

Our aim in 'Are There Neural Correlates of Consciousness?' was to call attention to some problematic assumptions of one widespread approach to investigating the relation between consciousness and the brain — the research programme based on trying to find neural correlates of the contents of consciousness (content-NCCs). Our aim was not to cast doubt on the importance of neuroscientific research on consciousness in general (contrary to **Baars's** impression). Nor was it to engage in philosophical debates far removed from the concerns of scientists (as **McLaughlin & Bartlett** may think). Rather, it was to target some problematic assumptions of a particular empirical research programme, and by bringing them to light, to suggest that there may be other, more profitable ways to investigate the contribution of brain processes to conscious experience than searching for content NCCs. Most of the commentators (**Bayne, Freeman, Hardcastle, Haynes & Rees, Hohwy & Frith, Metzinger, Myin, Roy, Searle, Van Gulick**), though certainly not all (**Baars, Jack & Prinz, McLaughlin & Bartlett**) seem to have read us this way, and we are grateful for their critical reflections on our article. In this Authors' Reply, we cannot respond in detail to every point raised by the commentators, so we shall limit ourselves to addressing the most important issues that we see arising from the commentaries collectively.

I: The NCC Research Programme

Most neuroscientists believe that experience happens in the brain. For every experience, it is assumed, there is a neural process whose occurrence is sufficient for that experience. Moreover, most scientists assume that if we could understand what is going on at the neural substrate of an experience, then we would understand how the brain's action produces states of consciousness.

Correspondence: Alva Noë, Department of Philosophy, University of California, Berkeley CA 94720-2390, USA. *Email: noe@socrates.berkeley.edu*
Evan Thompson, Philosophy Department, York University, 4700 Keele Street, North York, Ontario, M3J 1P3, Canada. *Email: evant@yorku.ca*

Journal of Consciousness Studies, 11, No. 1, 2004, pp. 87–98

Of course, right now no one has a clue how neural processes give rise to experience. So right now we wouldn't know a minimally sufficient neural substrate of experience even if it were staring us in the face. There is nothing controversial about this point. On the contrary, this very 'explanatory gap' motivates in part the NCC research programme. We don't have a clue how neural states produce experience, so let's look for neural processes that *correlate* with experience. Once we find them, we can then try to explain them in causal terms, 'in the hope that ... this will make the problem of qualia easier' (Crick and Koch, 2003, p. 119).

Is it reasonable to be so confident that we will one day be able to explain consciousness in terms of processes at the NCC? The NCC research programme is like a V8 engine in a car with no wheels — metaphysical power but no epistemic vehicle. The engine revs, the car trembles and roars, exhaust pours out, but where is the forward movement? Our main aim in our article was to call attention to this curious intellectual predicament. Scientists believe there must be an NCC, but given no theory of what an NCC is good for (no explanation of how an NCC could be a minimally sufficient neural substrate for experience), there's no good reason to take one neural process rather than another to be an NCC.¹ The NCC research programme, far from rescuing the problem of consciousness from the hands of philosophers (Crick, 1996, p. 486), is an act of metaphysical faith.

Research on neural correlates of consciousness is committed in practice to what we named the matching content doctrine. According to this doctrine, for every perceptual experience *E*, there is a neural representational system *N*, such that (i) *N* is sufficient for *E*; and (ii) there is a match between the content of *N* and the content of *E*. If *N* satisfies these conditions, then *N* is the NCC for content (the content-NCC) for *E*. We reject the matching content doctrine because we think there are no neural representational systems satisfying the conditions for being the content NCC for an experience. For this reason, we believe there are no content-NCCs. Our argument is straightforward: there can be no *match* between the content of neural representational systems and the content of experiences because neural content and experiential content are incommensurable (at least as neural content is usually understood in the NCC research programme, for instance as receptive-field or RF content.)

Some commentators worry that the matching content doctrine is a straw man (**Baars, Hohwy & Frith, McLaughlin & Bartlett, Jack & Prinz, Metzinger, Roy**). **McLaughlin & Bartlett**, for example, point out that many philosophers reject *intentionalism*, the idea that the phenomenal character of an experience is determined by its representational content. Hence to explain why an experience has the representational content it does in terms of the content of neural representational systems would still leave the phenomenal consciousness unexplained. Similarly, **Metzinger** accuses us of equivocating between the two basic notions

[1] For these reasons, we think **Hardcastle's** analogy with the weather is not apt. In the case of the weather, we understand many of the causal processes involved, but must abstract away from the microlevel detail in order to describe large-scale patterns. In the case of consciousness, we have no comparable understanding of the causal processes involved at either local or global neural levels. This lack of understanding is precisely the explanatory gap.

of intentional content and phenomenal content. We grant this point about intentionalism, which is not unfamiliar to us. The fact that many philosophers believe intentionalism is wrong, however, is not evidence that scientists working within the NCC research programme are not in fact guided by the matching content doctrine. Clearly they are. Crick and Koch write: ‘Whenever some information is represented in the NCC it is represented in consciousness’ (1998, p. 98). We can see the doctrine at work in studies of binocular rivalry, such as Tong *et al.*’s 1998 study using fMRI in humans. The authors exploited the well-known selectivity of fusiform gyrus to face stimuli and parahippocampal neurons to place stimuli. Using BOLD contrast signals, they established that when a face image is dominant in rivalry, FFA activity is high and PPA activity low, and that when the house image is dominant, PPA activity is high and FFA activity low. Indeed, they found that the levels of activation were the same as measured under nonrivalrous conditions, when stimuli were varied in such a way as to mimic binocular rivalry. On this basis, they write: ‘FFA and PPA activity is tightly linked to visual awareness during rivalry, reflecting both the *content* and duration of each percept’ (Tong *et al.*, 1998, p. 755, *our emphasis*).

The content-NCC reasoning is clear: (i) Certain neurons (in FFA) respond preferentially to the presence of faces in their receptive fields, whereas others (in PPA) respond to houses. (ii) In non-rivalrous conditions, these different groupings of cells respond when the subject is presented with a face-image or a house-image stimulus. (iii) The cells in FFA respond (more or less) equivalently in rivalrous conditions when the subject reports seeing a face, and they cease to respond when the subject reports seeing a house (even though the face-image is still present). And vice-versa. (iv) The subject has a perceptual experience as of a face when (and only when) these neurons with face receptive fields fire. (v) Conclusion: these neurons (in FFA and PPA) form (part of) the content-NCC for the experience of a face and house (respectively).

The reliance on the receptive content of neural systems is made explicit in work of Lumer, Friston and Rees (1998). In their study, they sought to see whether they could ‘provide an answer to the central issue of multistable perception — whether a specific machinery mediates the ongoing selection among sets of neuronal events competing for visual awareness’. They continue: ‘Because the rivalry and replay conditions yield similar perception and behavior, we expected them to engage common neural pathways associated with *the internal representation of visual scenes* and the generation of appropriate motor responses’ (pp. 1931–2, *our italics*). This work relies explicitly on the assumption that there are neural systems with a definite representational content. They seek to identify what further neural events or processes select among representations.

McLaughlin & Bartlett may be right that the discovery of content matches would leave the explanatory gap unbridged (for reasons having to do with the inadequacy of intentionalism). From the brain imaging trenches, however, such a philosophical nicety is apt to seem unimpressive. If we did find neural representational systems that both correlated with experience and matched experience in content, then we would certainly find it difficult to resist the conclusion that

those neural systems were at least part of the substrate of the experience. Overall, **McLaughlin & Bartlett** seem to miss the point of our target article. Our concern was to criticize a scientific working hypothesis. We do not aim to disprove this hypothesis, as it were logically. We grant that the minimal neural substrate thesis (even jointly with the isomorphism constraint) does not logically entail the matching content doctrine. Nevertheless, the matching content doctrine is a natural expression of these ideas, and one that plays an important role in neuroscientific research. It is clearly at work in the binocular rivalry research. It is singled out by Chalmers (2000) in his article on neural correlates of consciousness. These are the reasons why we criticize it, and why, in criticizing it, we cast critical light on the minimal neural substrate thesis.

II: Problems with the Matching Content Doctrine

The matching content doctrine faces a dilemma. On the one hand, RF-content is too thin to sustain a match with perceptual experience. On the other hand, a richer account of content threatens to be too rich to be the content of neural systems (as distinct, say, from the content of more abstract states).

In our article we claimed that perceptual content, unlike RF-content, is structurally coherent, intrinsically experiential, active and attentional. Most commentators accepted this characterization of perceptual experience, but a number of them challenged our claims that RF-content can be none of these things (**Bayne, Hardcastle, Haynes & Rees, Hohwy & Frith, Jack & Prinz, McLaughlin & Bartlett**). We now offer a simpler characterization of experiential content, and show why there can be no match with RF-content.

The content of perceptual experience has an important feature that RF-content does not and cannot have. Perceptual experience is perspectively self-conscious (Hurley, 1998; Noë, 2002; 2004a,b). A visual experience not only presents things as being a certain way; it presents how things are in relation to oneself. When you see, you see how things are from your vantage point. **Bayne** and **McLaughlin & Bartlett** point out correctly that perceptual experiences are not the only things that have this sort of perspectival representational content. Beliefs, propositions, photographs, and perhaps even RF-content can be perspectively representational in this way. Yet there is an important difference between these cases and perceptual experience. The content of perceptual experience is presented in such a way that to have the content is for one to know, implicitly and practically, that had things been different, they would have looked different, and that had one moved, and so changed one's spatial relation to the environment, things would have looked different. The difference between things looking thus and such to one, and one's merely thinking that they are thus and such, comes out in the presumed dependence of the content, in the perceptual case, on what is going on around one and on one's actual and possible movements. This dependence is part of the content of the experience itself.²

[2] A similar point was made by Searle (1983) in his account of the causal-indexical component of the intentional content of perception. According to Searle, when you see, it is not only the case that you

Neither belief, nor thought, nor photographs, nor sentences are dependent on the world and movement in this way. This thought is what we had in mind, in the article, when we said that perceptual content is intrinsically experiential. Our thought was that perception is not, as it were, a propositional attitude such that one can factor out the attitude and be left with a propositional content.

When we say that RF-content is too thin to sustain a match with the content of experience, we mean that RF-content lacks the sort of perspectival self-consciousness that is the hallmark of perceptual experience. The problem here is not that we do not understand this feature of experience, or that we cannot make sense of the attribution of this sort of content to an experiencer. The problem, rather, is that we do not understand how a neural representational system could have this sort of content. Indeed, it is difficult to see how anything other than a whole perceiver as an intentional agent could be the bearer of this sort of content. This difficulty is the second arm of the dilemma.

Let us emphasize that it is no part of our criticism of the matching content doctrine that we take NCC researchers to believe that the content of experience could be given by the RF-content of individual neurons, or that ‘consciousness can be located in individual neurons’, as **Jack & Prinz** and **Hardcastle** suggest. It is not at all clear to us why they should misunderstand our position in this way. That Jack and Prinz do misunderstand our argument comes out when they continue: ‘It is certainly plausible that every conscious experience corresponds to some population of neurons, however complex. The NCC research programme is trying to identify these populations to see whether they share any *signature features*. That exciting line of research is not threatened by the truism that experiences are rich, multimodal events’ (p. 54, our emphasis). As for the first point, who could object to the claim that there is likely to be a ‘correspondence’ between populations of neurons (‘however complex’) and experiences? Certainly we do not. But what do Jack and Prinz have in mind when they inquire whether the populations ‘share any signature features’? How would one identify such a feature if it presented itself? Our positive point is this: it is a mistake to think that the signature of the neural substrate will be a content match (see **Searle’s** point (1) for another line of argument in support of this point, as well as the helpful and insightful discussions by **Myin** and **van Gulick**).

III: Is Partial Content Match Enough?

Some commentators challenge our demand that there be a content match at the NCC (**Hohwy & Frith, McLaughlin & Bartlett, Roy**). **Hohwy & Frith**, for example, suggest that we set the bar on *correspondence of content* too high. If identity means ‘all the same content in the same representational format’, then there will never be interesting content matches. All that could match a

have a visual experience that depends causally on what the experience is of or about. In addition, it is part of the content of the experience that there is this very causal dependence. Although we would not accept this particular formulation (for reasons we need not go into here), we do endorse the basic insight: the perceiver and the world enter into the content of experience as background conditions, as it were. Perceptual experience is in this way intrinsically perspectivally-self-conscious.

photograph in content would be another photograph, and all that could match a sentence would be another sentence. This point is a fair one. But notice: the matching-content requirement is not *our* requirement. The point of our discussion was to bring to light the question: what degree of correspondence would be needed to justify the claim that one could explain the experience in terms of what is going on in the neurons?

Contrary to what **Hohwy & Frith** and others seem to think, partial agreement may not be enough. To see why, consider how the story would go. You have a visual experience of a vertical line. There is a neural representational system whose activation represents the presence of a vertical line, and whose activity correlates with your experience. Of course, the RF-content of that system does not match your experience, it only agrees with it partially. The experience, but not the neural system, for example, represents the line against a background. Nevertheless, as experimenters we can be happy with such a partial match, and then work to fill in the picture, finding neural representational systems elsewhere that represent other important features of the experience.

This is a reasonable story, but it is something of a disappointment.

First, the story is a concessive. What has become of the ‘signature features’ **Jack & Prinz** find so exciting? The account boils down to the admission that we are not going to find discrete systems whose RF-content matches the content of the experience. The minimal neural substrate of even the simplest experience is likely to be widely distributed spatially (and perhaps temporally) in the brain. Indeed, Rees, Kreiman and Koch (2002) seem to advocate precisely such a deflationary account of the NCC in a recent review article.

Second, we knew before we undertook the NCC research programme that various neural systems are necessary for conscious vision. Have we moved significantly beyond this familiar starting point?

Third, when all is said and done, the existence of distributed partial content agreements does nothing to address the issue about the supposed sufficiency of activity at the NCC for experience. In what way, if any, is the neural binding of distributed partial content agreements supposed to be sufficient for experience? Our claim in the article was that such activity may be necessary, but not sufficient for mental life. It is only the healthy brain of the healthy animal that supports consciousness.³

Fourth, at best, it would seem, we are left with the conclusion that causal interactions between brain and world place *constraints* on possible contents of experience. This conclusion is right, and important. It is also something we knew before we ever put a subject in an fMRI scanner or wrote up a grant proposal.

We want to be careful to dispel the impression that our main point here is negative. We believe the explanatory gap ought to be bridgeable (although not by the discovery of signature features of NCCs). Moreover, we believe that research into the neural grounds of experience (for example work on neural correlates of

[3] This is what we meant in our article when we wrote that neural processes causally enable but do not constitute our embodied mental life. We find it odd that **Hardcastle** thinks that ‘From a materialist perspective, this has to be false’ (p. 42).

binocular rivalry) is in fact bringing to light important information about the way neural activity subserves experience. We also accept **Hohwy and Frith's** suggestion that it is reasonable for work in this area to proceed in piecemeal fashion, step by step. In order for such research to bridge the explanatory gap, however, much more attention needs to be given to the sorts of questions about content match and agreement that we, and our commentators, are now addressing.

IV: Internalism and the NCC Research Programme

Some commentators challenge our arguments for an externalist conception of experience, i.e., for the idea that the causal basis of experience is not exclusively neural (**Bayne, Hohwy & Frith, Jack & Prinz, McLaughlin & Bartlett, Searle**). Their responses are somewhat surprising to us, because it was never our intention to advance arguments for externalism (or against internalism) in this paper. Rather, our aim was, as we wrote: 'to argue that the NCC research programme rests on a philosophically controversial internalist conception of the content of perceptual experience' (p. 20). We went on to write: '*We cannot show here that this conception is wrong*. But in showing that the NCC programme pre-judges an outstanding philosophical issue, we call into question Crick's all too confident claim that scientific research on the neural correlates of consciousness rescues the problem of consciousness from the hands of philosophers.... Because the internalist conception of conception *might* be wrong, the fate of the NCC research programme hangs on the resolution of a significant philosophical controversy' (emphasis added).

We will not attempt here to defend externalism about experience.⁴ We do wish, however, to correct one misunderstanding on the part of **Jack & Prinz**. They point out correctly that externalism entails at most that the NCC is the neural contribution to consciousness. The rhetorical force of this statement suggests that they think we might be inclined to dispute this point. They write that 'In the hands of Noë and Thompson, externalism becomes an eccentric doctrine that locates consciousness outside of the organism. To our minds, this is the kind of philosophical manoeuvring that prevents scientists from taking philosophers seriously. Rather than clarifying concepts, it obfuscates by conflating relational conditions on representation individuation with claims about their literal location' (p. 55).

It is *not* our view that consciousness is outside the head, but rather that some of the causal substrates of consciousness *might* be. Whether or not the *vehicles* of consciousness cross the skull boundary is an empirical question whose answer has not yet been decided, because we do not yet have an explanation of consciousness (see Thompson and Varela, 2001; Hurley and Noë, 2003; Noë, 2004a,c; Thompson, forthcoming). Neural states may not be sufficient for human and animal experience, but they are surely necessary (i.e., they surely make a contribution to the determination of consciousness).

[4] Each of us has his own views on this issue: see Noë (2004a,c); Hurley and Noë (2003); Thompson and Varela (2001); Thompson (forthcoming).

Jack & Prinz assert that externalism does not entail that consciousness is not in the head, and they cite Burge ('arch-externalist') to make their point: he grants that edge-detectors are literally in the head, arguing only that their individuation as edge-detectors depends on their relation to something outside the head (edges). Davidson (1987) made a similar point about sunburns (see also Wilson, 1995): They are literally *on* the skin, but they wouldn't *be* sunburns if not for a certain historical, causal relation that they bear to the sun itself. The point here is familiar: externalism about content of the sort advocated by Burge, Putnam, Dretske, Dennett and others is compatible with internalism about the vehicles of content (Hurley, 1998; Noë, 2001).

Be that as it may. Externalism about the *vehicles* of content (advocated in print by Hurley, 1998; Hurley and Noë, 2003; Rowlands, 2002; 2003; Clark and Chalmers, 1998; and Noë, 2004a,c), however, *does* entail that neural systems are not sufficient for consciousness. This view does not rule out the value of studying the neural basis of consciousness. But it does rule out the possibility that the NCC research programme can succeed.

Behind the purple prose is the ferocity of dogmatic conviction. **Jack & Prinz**, like many philosophers and scientists, *assume* that the causal substrates of experience are entirely in the head. Perhaps they are right. But the issue is far from settled yet.

V: The Binding Problem and Consciousness as a Unified Field

Searle observes that the binding problem remains even if, as we argued, the binding problem as it is standardly articulated is malformed because it presupposes an atomistic conception of content: 'There is ... more than one binding problem. One problem is: how does the brain create a unified conscious field?... Another form ... is how within the unified field do we get unified perceptual experiences of particular objects' (p. 80). These are important problems, but great care is required if we are to formulate them coherently. For example, it is a mistake to think that the mere fact that neural processes subserving the perception of colour, form and movement are spatially distributed raises a special problem about how those distributed features are bound together in the experience of a single moving object with a colour and form. Just as there is no reason to believe that neurons for colour must be coloured, so there is no reason to think that neurons for a spatially coherent object need to be spatially unified. In general, to investigate the unity of consciousness and the unity of objects of perception, we are inclined to place much more explanatory weight than Searle apparently would on the dynamic sensorimotor context of neural processes. For instance, our approach would seek to account for unified perceptual experiences of particular objects by appeal to invariant patterns of dynamic sensorimotor interdependencies between organism and environment, and not simply internal brain processes (O'Regan and Noë, 2001).

Searle also argues that the neural substrates of individual conscious perceptual states should not be thought of as sufficient for the occurrence of those states, for those states themselves presuppose the background consciousness of the animal.

Perceptual experiences are, as it were, modulations of a pre-existing conscious field. Here he draws from an article in which he distinguishes between ‘building block’ and ‘unified field’ approaches to consciousness (Searle, 2000). What Searle means by the building block approach, which he criticizes, is more or less equivalent to the content-NCC approach as we described it in our article. This approach aims to isolate neural correlates of the contents of consciousness for specific types of individual sensory experiences, such as the visual experience of a perceptually dominant stimulus in binocular rivalry. An underlying and problematic assumption of this approach is that consciousness is made up of various building blocks corresponding to constituent individual experiences, which are then somehow bound together to constitute the unity (or apparent unity) of consciousness. As Searle points out: ‘Given that a subject is conscious, his consciousness will be modified by having a visual experience, but it does not follow that the consciousness is made up of various building blocks of which the visual experience is just one’ (2000, p. 572). According to the unified field approach, on the other hand, an individual experience or conscious state (such as the visual perception of a face) is not a constituent (in the building block sense) of some aggregate conscious state, but is rather a modification within the field of a basal or background consciousness: ‘Conscious experiences come in unified fields. In order to have a visual experience, a subject has to be conscious already, and the experience is a modification of the field’ (ibid.). This approach accordingly focuses its investigations on the neurophysiological processes that characterize the dynamic modulation of the conscious field, and the differences in large-scale dynamic brain activity across various conscious and unconscious conditions (Varela, 1995; Tononi and Edelman, 1998; Freeman, 1999; Varela *et al.*, 2001). **Freeman** has been one of the main pioneers of this sort of approach, and he briefly describes a few of its features in his commentary. We are very sympathetic to this approach, especially when placed in the context of an embodied approach to the mind, and one of us has written specifically in this area (Thompson and Varela, 2001; Varela and Thompson, 2003; Lutz and Thompson, 2003).

VI: Some Unfinished Business

In this section we address two issues raised in commentaries that fall outside the range of issues we have discussed so far.

Conceptual versus nonconceptual content

McLaughlin & Bartlett criticize us for failing to attend to the distinction between the conceptual and the nonconceptual content of perceptual experience (and between an associated distinction between phenomenal and nonphenomenal senses of ‘looks’). They claim that the matching content doctrine is most plausible for the nonconceptual content of experience. They voice a suspicion that we are not ‘attuned’ to the relevant distinctions, and that this insensitivity on our part explains mistakes in our analysis of the visual experience of the presence of partially occluded parts of a cat, and also the claim that egocentric space is an

aspect of the content of visual experience. They also point out that although the representational content of neural systems may fail to be subject to constraints of holism and normativity, this is also likely to be true of nonconceptual content. In addition, they claim that nonconceptual content, like the content of neural systems, can be personal-level, even though it is not subject to constraints of holism and normativity (for it figures in the explanation of voluntary behaviour).

We appreciate the interest of these topics and one of us has explored them at length in print (Noë, 1999; 2002; 2004a). We also grant that we ought to have brought this issue out in the open in our discussion. Most scientists, however, and many philosophers (including one of us), find debates in this area rather ‘scholastic’; and this led us to steer clear of it in the text. For now, let us briefly clarify our position.

The spatial content of visual experience *is* egocentric. Or so we would argue on the basis of first-person reflection, phenomenological analysis, and the theory of content. If **McLaughlin & Bartlett** are right that the egocentric aspect of visual content includes features outside nonconceptual content (such as, e.g., the conceptual content of perceptual expectations), then we are inclined to hold that we ought to give up the idea that the spatial content of visual experience is or could be nonconceptual. Alternatively, it may be that the relevant expectations thanks to the possession of which it is possible for the perceiver to have visual experiences with egocentric spatial content are themselves nonconceptual. Noë (2004a) has argued that the relevant skills are sensorimotor, rather than conceptual (although he also argues that, in important respects, sensorimotor skills are quasi-conceptual).

Similar points go for the perceptual presence of partially occluded objects. Surely when a dog sees a squirrel running up the side of a tree, it has a visual experience whose (nonconceptual) content is that of a squirrel running up the side of the tree. It is the running of a squirrel up a tree that would make that visual experience of the dog veridical (among other possible conditions). But now let us ask: does the dog experience the three-dimensionality of the tree? Or the existence of the tree in those places that are blocked from view by the squirrel’s body? Is the ability of perceptual experience to represent features of *that* sort a concept-dependent ability? If so, then we had better admit that perceptual experience doesn’t have nonconceptual content and that dogs have concepts.

These issues require further discussion than we cannot enter into here. Of particular importance is the interplay between the distinction between conceptual and nonconceptual content, on the one hand, and representational and nonrepresentational properties of experience, on the other.

Baars’ attack

Baars’ commentary is flagrantly unsympathetic and shows almost no evidence of even a cursory reading of our article. In contrast to the other scientists who contributed constructive critical commentaries, Baars distorts and misinterprets our discussion at every step. It is tempting to leave matters at that, but for the sake of the record we feel obliged to spell out the following points: (1) We never

‘cast doubt’ on ‘such basic facts’ as ‘that sensory consciousness requires brain processes to represent the world’ (p. 29). (2) We do not ‘set an unrealistically high standard’ for what neuroscientists are doing; we point to a number of problematic assumptions about content that the content-NCC research programme needs to address — which is precisely what **Haynes & Rees** and **Hohwy & Frith** do in their useful commentaries. (3) Nowhere in our article do we ‘attack scientific studies of consciousness’ (p. 29); we criticize certain assumptions of one particular research programme. (4) Our article is not a piece of ‘technical analysis’ in Baars’s idiosyncratic sense (p. 30). Some of our arguments are phenomenological, and we refer to writings in both analytical and phenomenological traditions. (5) We make no claims at all about ‘the essential function’ of consciousness; in particular we make no claim that ‘the essential function of consciousness is to match personal experience to the world’ (p. 30) (which, taken literally, seems barely coherent, for it implies that consciousness has the ‘essential function’ of matching itself to the world). (6) Baars accuses us of scrambling the evidence and its interpretation (p. 29), but consider his description of the Logothetis *et al.* results: ‘Logothetis and co-workers compared *conscious* and *unconscious visual input* to the two eyes of a macaque monkey’ (p. 30, our emphasis). This is confused. The input to the eyes — the stimulus — is neither conscious nor unconscious. Baars also states that ‘no cells in the object recognition areas (area IT/TE) responded to the unconscious input [sic]. All of them fired to the conscious stream [sic]’ (pp. 30–1). This is inaccurate. As we say in our article (p. 10), it was found that 90% of the recorded neurons in IT and STS fired when the monkey reported seeing the cell’s preferred pattern. (7) Baars states that we ‘claim there is no topographical mapping of the sensory surround in the conscious brain’ (p. 31). We make no such claim. (8) Finally, Baars seems to imply that we think consciousness does not depend causally on the brain (see his final two paragraphs), despite our saying in the very first paragraph of our article that we assume ‘states of consciousness causally depend on states of the brain’.

VI: Final Thoughts

For there to be progress in understanding the brain basis of consciousness, we need to advance on two fronts. First, we need better, more reliable characterizations of the phenomena of consciousness we seek to explain. As an example, we need accounts that allow us to focus (as we do above) on the perspectival self-consciousness of perceptual content. Second, we need better models of the content and function of neural systems themselves. The burden of our article was to say why we think it is unlikely we can make significant progress in these areas by following the content-NCC approach. Several commentators criticized us for not proposing a constructive alternative: as **Freeman** puts it, we left underdeveloped our preferred alternative to the content-NCC approach (p. 38). We cannot develop such an alternative here, but we would like to close by noting that each of us is engaged in the development of alternative models, ones that emphasize the dynamic sensorimotor context of neural activity, the need for better

phenomenological accounts of conscious experience, and the need for large-scale, integrative models of neural dynamics.⁵

References

- Chalmers, D.J. (2000), 'What is a neural correlate of consciousness?', in *Neural Correlates of Consciousness*, ed. Thomas Metzinger (Cambridge, MA: The MIT Press/A Bradford Book).
- Clark, A. and Chalmers, D.J. (1998), 'The extended mind', *Analysis*, **58**, pp. 10–23.
- Crick, F. (1996), 'Visual perception: Rivalry and consciousness', *Nature*, **379**, pp. 485–6.
- Crick, F. and Koch, C. (1998), 'Consciousness and neuroscience', *Cerebral Cortex*, **8**, pp. 97–107.
- Crick, F. and Koch, C. (2003), 'A framework for consciousness', *Nature Neuroscience*, **6**(2), pp. 119–26.
- Davidson, D. (1987), 'Knowing one's own mind', *Proceedings and Addresses of the American Philosophical Association*, **60**, pp. 441–58.
- Freeman, W. (1999), *How Brains Make Up Their Minds* (London: Nicolson & Weidenfeld).
- Hurley, S.L. (1998), *Consciousness in Action* (Cambridge, MA: Harvard University Press).
- Hurley, S.L. and Noë, A. (2003), 'Neural plasticity and consciousness', *Biology and Philosophy*, **18**, pp. 131–68.
- Lumer, E.D., Friston, K.J. and Rees, G. (1998), 'Neural correlates of perceptual rivalry in the human brain', *Science*, **280**, pp. 1930–3.
- Lutz, A. and Thompson, E. (2003), 'Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness', *Journal of Consciousness Studies*, **10** (9–10), pp. 31–52.
- Noë, A. (1999), 'Thought and experience', *American Philosophical Quarterly*, **36** (3), pp. 257–65.
- Noë, A. (2001), 'Experience and the active mind', *Synthese*, **129** (1), pp. 41–60.
- Noë, A. (2002), 'Is perspectival self-consciousness nonconceptual?', *The Philosophical Quarterly*, **52** (207), April, pp. 185–94.
- Noë, A. (2004a), *Action in Perception* (Cambridge, MA: The MIT Press).
- Noë, A. (2004b), 'The enactive approach to perception', Interdiscipline online symposium.
- Noë, A. (2004c), 'Experience without the head' in *Perceptual Experience*, ed. Tamar Gendler and John Hawthorne (Oxford: Oxford University Press).
- O'Regan, J.K. and Noë, A. (2001), 'A sensorimotor account of vision and visual consciousness', *Behavioral and Brain Sciences*, **25** (4), pp. 883–975.
- Rees, G., Krieman, G., and Koch, C. (2002), 'Neural correlates of consciousness in humans', *Nature Reviews Neuroscience*, **3**, pp. 261–70.
- Rowlands, M. (2002), 'Two dogmas of consciousness', *Journal of Consciousness Studies*, **9** (5–6), pp. 158–80.
- Rowlands, M. (2003), *Externalism: Putting Mind and World Back Together Again* (Chesham: Acumen).
- Searle, J. (1983), *Intentionality: An Essay in the Philosophy of Mind* (Cambridge: Cambridge University Press).
- Searle, J. (2000), 'Consciousness', *Annual Review of Neuroscience*, **23**, pp. 557–78.
- Thompson, E. (forthcoming), *Radical Embodiment: The Lived Body in Biology, Cognitive Science, and Human Experience* (Cambridge, MA: Harvard University Press).
- Thompson, E. and Varela, F.J. (2001), 'Radical embodiment: Neural dynamics and conscious experience', *Trends in Cognitive Sciences*, **5**, pp. 418–25.
- Tong, F., Nakayama, K., Vaughan, J. and Kanwisher, N. (1998), 'Binocular rivalry and visual awareness in human extrastriate cortex', *Neuron*, **21**, pp. 753–59.
- Tononi, G. and Edelman, G. (1988), 'Consciousness and complexity', *Science*, **282**, pp. 1846–51.
- Wilson, R.A. (1995) *Cartesian Psychology and Physical Minds: Individualism and the Science of the Mind* (Cambridge: Cambridge University Press).
- Varela, F.J. (1995), 'Resonant cell assemblies: A new approach to cognitive functions and neuronal synchrony', *Biological Research*, **28**, pp. 81–95.
- Varela, F.J. and Thompson, E. (2003), 'Neural synchrony and the unity of mind: A neurophenomenological perspective', in *The Unity of Consciousness*, ed. A. Cleeremans (New York: Oxford University Press).
- Varela, F.J., Lachaux, J.-P., Rodriguez, E. and Martinerie, J. (2001), 'The brainweb: Phase synchronization and large-scale integration', *Nature Reviews Neuroscience*, **2**, pp. 229–39.

[5] On the dynamic sensorimotor approach see Hurley and Noë (2003) and Noë (2004a); on the integration of phenomenology and neurodynamics see Lutz and Thompson (2003).