SENSORIMOTOR SKILLS AND PERCEPTION

by Andy Clark and Naomi Eilan

I—Andy Clark

COGNITIVE COMPLEXITY AND THE SENSORIMOTOR FRONTIER

ABSTRACT What is the relation between perceptual experience and the suite of sensorimotor skills that enable us to act in the very world we perceive? The relation, according to 'sensorimotor models' (O'Regan and Noë 2001, Noë 2004) is tight indeed. Perceptual experience, on these accounts, is *enacted* via skilled sensorimotor activity, and gains its content and character courtesy of our knowledge of the relations between (typically) movement and sensory stimulation. I shall argue that this formulation is *too* extreme, and that it fails to accommodate the substantial firewalls, dis-integrations, and special-purpose streamings that form the massed strata of human cognition. In particular, such strong sensorimotor models threaten to obscure the computationally potent *insensitivity* of key information-processing events to the full subtleties of embodied cycles of sensing and moving.

I

The Painter and the Perceiver. Seeing, according to Noë 2004, is like painting. Painting is an ongoing process in which the eye probes the scene, then flicks back to the canvas, then back to the scene, and so on, in a dense cycle of active exploration and partial, iterated cognitive uptake. It is this cycle of situated, world-engaging activity that constitutes the act of painting. Seeing (and more generally, perceiving) is likewise constituted (Noë claims) by a process of active exploration in which the sense organs repeatedly probe the world, delivering partial and restricted information on a need-to-know basis. It is this cycle, of situated, world-engaging, whole animal activity, that is the locus, on Noë's account, of genuine cognitive interest, at least for perceptual experience.

An important implication of this, according to Noë, is that appeals to internal representations (if such there be) cannot tell the whole story, either for painting or for seeing:

> The causally sufficient substrate of the production of the picture is surely not the internal states of the painter, but rather the dynamic pattern of engagement among the painter, the scene and the canvass. Why not say the same thing about seeing? Seeing, on this approach, would depend on brain, body and world. (Noë 2004, p. 223)

In the case of seeing (and perceiving in general) a theoretical construct peculiarly well-suited to this dynamic target is the notion of sensorimotor dependencies (aka sensorimotor contingencies—see O'Regan and Noë 2001). Sensorimotor dependencies are relations between movement or change and sensory stimulation. Such relations may be of many kinds (see below) but what they all have in common is that they concern a kind of loop or cycle linking real-world objects and properties with systematically changing patterns of sensory stimulation. These changing patterns of sensory stimulation may be caused by the movements of the subject (this is the central case), as when we use head and eye movements to scan a visual scene. Or they may be caused by movements of the object itself, or be due to other elements in the environmental frame (such as changes in illumination or light source). In addition, some features of these various kinds of changing pattern will be due to properties of the objects themselves (for example, the self-similarity of a straight horizontal line along its length, giving rise to an unchanging pattern of retinal stimulation as the eye tracks along the line (see O'Regan and Noë 2001, p. 942), while others will be due to the idiosyncrasies of the human visual apparatus (for instance, the same straight line projected onto the retina distorts dramatically as the eye moves up and down, due to the curvature of the eyeball (op. cit., p. 941)).

Usually, many kinds of sensorimotor dependence (some, as we have just seen, much more truly sensorimotor than others) are in play when we see an object. It is our implicit knowledge of these sensorimotor dependencies that explains, according to the strong sensorimotor model, both the contents and the character (visual, tactile, auditory, etc.) of our perceptual experiences. This stress on knowledge of (or expectations concerning) sensorimotor dependencies is meant as an alternative to standard appeals to qualia conceived as intrinsic, 'sensational' properties of experience. Instead of appreciating such mysterious intrinsic qualities in experience, it is suggested, we *enact* (that is, by acting bring into being) perceptual experience.[1] In the case of shape and spatial properties, for example:

> [T]he enactive view denies that we represent spatial properties in perception by correlating them with kinds of sensation. There is no *sensation* of roundness or distance, whether tactile, visual or otherwise. When we experience something as a cube in perception, we do so because we recognize that its appearance varies (or would vary) as a result of movement, that it exhibits a specific sensorimotor profile. (Noë 2004, pp. 101–2)

Noë's account, taken as a whole, has at least three apparent virtues, some key aspects of which I shall be seeking (as far as possible) to preserve.

First, and most importantly, there is the emphasis on skills rather than on qualia as traditionally conceived.[2] Skill-based accounts (see also Pettit 2003, Clark 2000, Matthen 2005, and of course Dennett 1991) offer a powerful antidote to the venom of zombie thought experiments.[3] In particular, the strong sensorimotor account would—if all worked out according to plan—ensure that sameness of world-engaging sensorimotor skills and discriminatory capacities implied sameness of perceptual experience. More demonstrably, the emphasis on world-engaging loops and knowledge of sensorimotor dependencies affords an elegant and compelling account of a range of real-world phenomena involving sensory substitution and neural re-wiring.

The classic example here is Tactile-Visual Sensory Substitution (TVSS; see Bach-y-Rita and Kercel 2002 for a recent review). Equally impressive, though perhaps less well known, is the

1. This general claim most strongly characterizes the work of O'Regan and Noë 2001 and Noë 2004. Historically, the view has roots that span science (especially ecological psychology; see Gibson 1979) and several influential philosophical traditions (ranging from Husserl 1907, Heidegger [1927] 1961 and Merleau-Ponty [1945] 1962, to Ryle [1949] 1990, MacKay 1967, and on to the original enactive approach of Varela, Thompson and Rosch 1991). It is also consistent with (but goes far beyond) the project of understanding mind and cognition in ways that are heavily 'action-oriented' (Clark 1997) and that stress the importance of body, action, and the canny use of environmental structure (for example, Hurley 1998, Ballard et al. 1997, Hutchins 1995, Churchland, Ramachandran and Sejnowski 1994, Thelen and Smith 1994)).

2. For an excellent (though itself sceptical) account of the traditional conception of qualia, see Pettit 2003.

3. For Noë's own take on such thought experiments, see, for example, Noë 2004, p. 124.

auditory-visual substitution system (discussed in some detail in O'Regan and Noë 2001), known as The Voice (see Meijer 1992). In this system, visual inputs to a head-mounted camera are systematically translated into audible patterns. Objects high in the visual field yield high-pitch sounds, while low ones yield low-pitch sounds. Lateral location is indicated by the balance of stereo sound, brightness by loudness of sound, and so on. Crucially, as you move the camera around, the sound changes, and over time subjects begin to learn the signature patterns (the sensorimotor dependencies) characteristic of different objects. In the original versions, subjects learnt to distinguish plants from statues, crosses from circles, and such like.

The overall effect, though powerful, fell short of creating a truly visual experience. But the claim of sensorimotor dependency theory is bold and clear: to whatever extent it is *possible* to recreate the same body of sensorimotor dependencies using an alternative route, you will recreate the full content and character of the original perceptual experience. This explains, according to O'Regan and Noë, why some of Bach-y-Rita's subjects report, for example, feeling as if they were *seeing* a looming ball when fitted with a TVSS system. By stressing similarities and differences in the profile of sensorimotor dependencies, Noë-style accounts neatly explain both the sense in which such systems create quasi-visual experiences and the ways in which the experiences thus generated (currently) fall short of those supported by the original routes. For example, there is a clear sensorimotor signature for a looming object whose invariant characteristics are as well captured by patterns of sound or tactile stimulation as they are by the (more typical) patterns of retinal stimulation. Very fine-grained colour information, by contrast, is currently not well captured by these kinds of substitution system. In each case, however, what is at issue is not the presence or absence of mysterious, ineffable qualia, but simply the presence or absence of distinctive loops linking real-world objects and properties to changing patterns of sensory stimulation.

The same story explains, we are told, the remarkable results concerning the rewiring of visual inputs to auditory cortex in young ferrets (Sur et al. 1999). Here, 'auditory' cortical areas become (thanks to the early re-wiring) involved in the kinds of sensorimotor loop characteristic of vision, and appear to support fully normal visual capacities in the modified ferret:

> Appropriately embedded in a 'visual' sensorimotor dynamic, neural activity in 'auditory' cortex in young ferrets takes on 'visual' functions. (Noë 2004. p. 227)

In short, appeals to the shape of a space of signature sensorimotor dependencies here replaces appeal either to the intrinsic properties of sensations (qualia) or to their more hard-nosed (but arguably equally unexplanatory) cousins, the putative special properties of specific neural regions.

The sensorimotor account is thus meant to be successful in cases 'where neural accounts alone are explanatorily afloat' (Noë 2004, p. 226). What does the work is simply 'the way neural systems subserve the activity of the embodied and embedded animal' (op. cit.). For Noë, then, experience is

> not caused by and realized in the brain, although it depends causally on the brain. Experience is realized in the active life of the skilful animal. (Noë 2004, p. 226)

I shall argue that neural accounts need not be seen as quite so 'explanatorily afloat' even if we agree (as I think we should) that *certain* bodies of skill provide the key to understanding perceptual experience.

Second (in the list of possible virtues) is the sensorimotor model's recognition of the importance, power and scope of what (in the artificial neural network community) is known as prediction learning. Prediction learning is an ecologically plausible form of supervised learning. In supervised learning, an agent is provided with detailed feedback concerning the desired output for a given input. Since such training seems to require a well-informed, continually present teacher, its ecological plausibility, for most real-world learning situations, looks doubtful. In some cases, however, the world itself provides, at the very next time-step, precisely the training information we need. Such is the case if, for example, the task (typically, as presented to a Simple Recurrent Neural Net; see, for example, Elman 1995) is to predict the next sensory input itself, whether it be the next word in a sentence or the next 'frame' in an evolving visual scene. Such prediction, for a mobile embodied agent, often requires a double input, namely information concerning the current sensory state and information (for instance, in the form of efferent copy) concerning any motion command currently in play. Given these items of information, a prediction can be made concerning the likely next sensory state. Such a prediction, in the visual case, will thus need to take into account both features of the scene and any motions of the agent, and can immediately be tested against the actual sensory stimulations duly delivered by the world.

Prediction learning has shown itself to be a valuable tool for the extraction of a number of important regularities, such as those characteristic of grammatical sentences, of shape, and of object-permanence. In a sense, Noë and his collaborators are extending this proven paradigm to attempt to account for the full spectrum of perceptual experience, whose contents and character are said to be sensitively dependent upon acquired expectations (implicit knowledge) concerning the ways sensory stimulation will morph and evolve with movement and other kinds of input-altering change. This is, it seems to me, precisely the kind of knowledge that would be embodied in the weights and connections of a neural network trained using a prediction learning regime.

Prediction learning is computationally potent, demonstrably possible, and almost certainly biologically actual. The standard models are, however (as we just saw), resolutely sub-personal, with the predictions defined over sensory patterns that obtain outwith any conscious awareness. On Noë's account, however, a critically important sub-class of cases are defined over *consciously experienced perspectival properties* ('P-properties'; see Noë 2004, p. 83) of objects. These are depicted as objective but relational properties: properties belonging to a perceiver-object pair situated in some larger environment:

> That a plate has a given P-shape is a fact about the plate's shape, one determined by the plate's relation to the location of a perceiver, and to the ambient light. (Noë 2004, p. 83)

Importantly, P-properties are *also* depicted as 'looks of things, their visual appearances' (op. cit., p. 84), and thus as able to participate in *phenomenologically salient* bouts of prediction learning. Thus:

> To see a circular plate from an angle, for example, is to see something with an elliptical P-shape, and it is to understand how that perspectival shape would vary as a function of one's (possible or actual) movements . . . (Noë 2004, p. 84)

But whilst agreeing that prediction learning is a powerful *knowledge-extraction tool*, especially in the perceptual arena, I am not convinced that mature perceptual experience is then constituted by[4] the running of (what might be thought of as) the prediction software itself. That is to say, I am not convinced that appeal to predictions (or expectations) concerning the next sensory stimulation directly and exhaustively explains (sub-personally) or even characterizes (personally) perceptual experience.

We shall return to these issues in subsequent sections. For the moment, it is useful simply to distinguish three different questions that may be asked:

(1) What kinds of unconscious know-how drive or power our fluid sensorimotor engagements with the world?
(2) What do we implicitly know about how our conscious perceptual experience will vary during with movement or change?
(3) What determines the content and character of our conscious perceptual experience itself?

These questions are all different, but the strong sensorimotor model tends to offer the same kind of answer (one that invokes implicit knowledge of sensorimotor dependencies) to them all. I shall argue, however, that while the appeal to knowledge of sensorimotor dependencies might well be crucial to answering the first (as when an agent deploys 'emulator circuitry' to anticipate sensory input and hence drive smooth reaching, etc.[5]), it is by no means obvious what role it should play in the other two. Probably we do (regarding question (2)) have *expectations* concerning the ways conscious experience will alter as we move etc., but it is not obvious that these expectations are crucial to the experience itself. Indeed (and moving on to question (3)) there is considerable evidence that perceptual experience is linked to specific forms of neural processing that are systematically insensitive to much of the fine detail of the sensorimotor loops themselves, thus casting doubt on the strong sensorimotor response to both these questions.

4. Noë uses 'constituted by' to mean something like 'realized by', and that is the sense intended here.
5. For one worked-out account of this, see Clark and Grush 1999.

The third and final virtue I want (very briefly) to mention is rather general, but both important and surprisingly delicate. It is that the sensorimotor model is well poised to accommodate *narcissistic experience in an objective world*. Talk of cognitive agents that, by their own activity, 'bring forth their worlds' can seem mysterious, if not mystical (see Clark and Mandik 2002 for some discussion). But, by linking the contents and character of perceptual experience rather directly to acquired expectations concerning patterns of sensorimotor dependence, the enactive framework is able to do justice both to the notion of an objective, mind-independent reality and to the sense in which the world as perceived is the world of a specific type of embodied agent. Such a (perceived) world is characterized by a suite of distinctive sensorimotor dependencies, whose nature sensitively determines the way the world is experienced through the senses.

According to this account, differently embodied beings will not be able to directly experience our perceptual world, not because it is populated by its own mysterious qualia but because they lack the requisite 'sensorimotor tuning' (Noë 2004, p. 156). It is a virtue of the sensorimotor model that it allows us to address this thorny topic in a straightforward manner. But it is a vice (or so I shall later suggest) that in doing so it implies that differently embodied beings *necessarily* inhabit different 'perceptual worlds'.

In general, then, I shall argue that in several domains the strong sensorimotor model takes us one step too far. By stressing skills, abilities and expectations, such accounts begin to offer a genuine alternative to traditional qualia-based approaches to perception and perceptual experience. But by focusing so much attention upon the sensorimotor frontier, they deprive us of the resources needed to construct a more nuanced and multi-layered model of perceptual experience, and risk obscuring some of the true complexity of our own cognitive condition.

II

Sensorimotor (Hyper) Sensitivity. Strong sensorimotor models suffer, it seems to me, from a form of sensorimotor hypersensitivity. Such models, or so I shall argue, are hypersensitive to the very fine details of bodily form and dynamics and, as a result, are prematurely committed on a variety of prima facie open (empirical) questions concerning the tightness of the relation between perceptual experience and embodied action.

To begin to bring this rather general concern into focus, consider first the matter of what Clark and Toribio (2001) dub 'sensorimotor chauvinism'. A sensorimotor chauvinist, as we use the term, is someone who holds, without compelling reason, that absolute sameness of perceptual experience requires absolute sameness of fine-grained sensorimotor profile. Noë (2004) is clear enough about this commitment. For example, in a discussion of the extent to which TVSS systems support 'similarity of experience' (to normal vision) Noë asserts that:

> Tactile vision is vision-like to the extent that there exists a sensorimotor isomorphism between vision and tactile vision. But tactile vision is unlike vision precisely to the extent that this sensorimotor isomorphism fails to obtain. It will fail to obtain, in general, whenever the two candidate realizing systems differ ... in their ability to subserve patterns of sensorimotor dependence. (Noë 2004, p. 27)

Expanding on this idea, Noë adds that:

> Only a vibrator array with something like the functional multiplicity of the retina could support genuine (full-fledged, normal) vision. To make tactile vision more fully visual, then, we need to make the physical system on which it depends more like the human visual system. (Noë op. cit., pp. 27–8)

Despite the superficially liberal appeal, in these quotes, to 'functional multiplicity', the required identity (for precise sameness of experience) actually reaches far down into the structure of the physical apparatus itself, and demands very fine-grained similarities of body and gross sensory equipment. O'Regan and Noë are more explicit:

> For two systems to have the same knowledge of sensorimotor contingencies *all the way down* they will have to have bodies that are identical *all the way down* (at least in relevant respects). For only bodies that are alike in low-level detail can be functionally alike in the relevant ways. (O'Regan and Noë 2001, p. 1015)

While later on in Noë's single-author treatment he asserts that:

> Creatures with bodies like ours would have systems that are visual in the way ours are. Indeed, *only such systems can participate in the identical range of sensorimotor interactions that we participate in.* (Noë 2004, p. 159, my italics)

The position is thus that while some coarse-grained isomorphisms may be sufficient to begin to render the experience of a differently embodied being visual, the full glory of normal human visual experience depends upon a gross sensorimotor profile that very sensitively tracks the fine details of human embodiment (including, we saw, such matters as the precise curve of the eyeball).[6] Of course, even such a strong view need not be (as Noë (2004, p. 28) rightly points out) chauvinistic *if* the requirement of full sensorimotor isomorphism (for identity of experience) flows from a compelling theoretical model.

But does it? The claim in question (let's call it the claim of Fine-Grained Sensorimotor Dependence) is that every difference in fine-grained patterns of sensorimotor dependence will potentially impact any associated perceptual experience. Notice that this consequence does not in any way follow from the fact (if it is a fact) that prediction learning plays a key role in the *acquisition* of certain kinds of perceptual knowledge and understanding. For the upshot of such learning might well be forms of understanding that are systematically insensitive to some changes in sensory stimulation, while exaggerating others.

Notice also that the patterns of sensorimotor dependence in question cannot *themselves* be patterns in experiential space (in the space of appearances) on pain of triviality. For of course, every difference in experience implies some difference in experience. But if we step outside the phenomenological arena, then the claim (of Fine-Grained Sensorimotor Dependence) looks to involve the premature settling of what should be an open empirical question.

Thus suppose, to imagine a concrete case, that certain patterns of sensorimotor dependence concern the relations between movement and retinal stimulation. And suppose that some very small difference in embodiment makes a very small difference to such patterns. It is surely an open empirical question whether every difference in respect of such stimulation *makes a difference* to the content and character of any conscious perceptual experience

6. Thus we read that '... it turns out that there is good reason to believe that the sensorimotor dependencies are themselves determined by low-level details of the physical systems on which our sensory systems depend. The eye and the visual parts of the brain form a most subtle instrument indeed, and thanks to this instrument, sensory stimulation varies in response to movement in precise ways. To see as we do, you must have a sensory organ and a body like ours' (Noë 2004, p. 112).

that ensues. And the same will be true wherever in the processing story we choose to focus, even if we opt for patterns of cortical rather than retinal stimulation.

Systematic insensitivities might, in fact, serve some functional purpose. It is easy to imagine design and engineering considerations that would favour various kinds of buffering, filtering, and re-coding of perceptual inputs such that the contents and character of conscious perceptual experience might be determined at some considerable remove from the fine-grained details of sensorimotor loops. As we shall later see, there is some reason to believe that human perceptual experience is indeed determined at just such a remove, and that it involves tweaked and optimized representations that do not march sensitively in step with the flow of gross sensory stimulation.

It might be objected that the kind of hypersensitivity I am contesting is simply the price one pays for appealing to embodied skills as an alternative to traditional appeals to qualia. But this is not so. For the skills to which such deflationary accounts (among which I count the strong sensorimotor theory) appeal may *themselves* be coarse- or fine-grained, and may thus involve activities and capacities that are systematically insensitive to some of the goings-on at the sensorimotor frontier. For example, they may focus on what Matthen (2005) (and see also Pettit 2003) calls 'epistemic' skills: skills of sifting, sorting, classifying, selecting, choosing, re-identifying and comparing. These skills (which must, in any deflationary context, be said to constitute, rather than to call upon, perceptual experience) may depend on modes of processing and forms of internal representation that ultimately float free of the full spectrum of fine sensorimotor detail. Nor, finally, need the appeal to skills (rather than qualia) force us to abandon the notion of a distinctive personal level at which a cognitive agent has access to (some) information. That is to say, it should not force us to abandon the notion of that which is in some important sense *manifest* (see Pettit 2003) to the agent concerned.

I suspect that in his (admirable) eagerness to avoid the qualia trap, Noë has been led to define appearances rather too directly in terms of objective relations between objects and perceivers, with the result that whatever impacts this objective relation (more precisely, whatever impacts the way this relation unfolds during sensorimotor activity) is said to impact (if only very subtly) how things look to the agent. Other ways of unpacking a skill-based account need not, as we'll see, buy into this kind of picture. But before exploring such a possibility, it helps to introduce a missing layer of complexity in (at least some versions of) the strong sensorimotor model itself.

III

What Reaching Teaches. The complexity in question concerns the role of reason and planning in (what might be thought of as) the architecture of perceptual experience. According to O'Regan and Noë we are conscious of a specific visually presented state of affairs only when our practical knowledge about the ways movement will yield sensory change is actively invoked in the service of reason, planning and judgement. In such cases we do not merely exercise our mastery of sensorimotor contingencies, for we do this even when we are unaware of our own actions, as when returning a fast tennis serve or absent-mindedly driving along a familiar road. Rather, conscious awareness enters the scene when we make use of *that very same* knowledge of sensorimotor contingencies 'for the purposes of thought and planning' (O'Regan and Noë 2001, p. 944). On this account to consciously see is

> to explore one's environment in a way that is mediated by one's mastery of sensorimotor contingencies, and to be making use of this mastery in one's planning, reasoning and speech behavior. (O'Regan and Noë op. cit., p. 944)

The point of adding such a requirement is clear. Very often, when we exercise our implicit knowledge of patterns of sensorimotor dependence, no corresponding perceptual awareness ensues. To explain the difference, O'Regan and Noë invoke use in reason, planning and speech behaviour as a kind of spotlight that allows some (but not all) of our active knowledge of sensorimotor contingencies to condition perceptual awareness.[7]

7. Interestingly, this requirement (which is made much of by O'Regan and Noë) is nowhere in evidence in Noë's 2004 solo treatment. What we find there is just the bare idea of the active use of specific bodies of knowledge concerning sensorimotor dependencies in the guidance of behaviour. Noë (personal communication) picks this issue out as one where his views are in a state of flux. The guiding thought, he writes, is that 'being conscious of a feature is actively probing it—it's reaching out

There is, however, another possibility hereabouts, one that has significant empirical support and that is ultimately (or so I shall argue) at odds with the strong sensorimotor model. This is the possibility (Milner and Goodale 1995, Clark 2001, Jacob and Jeannerod 2003) that the contents of conscious perceptual experience are determined by the activation of a distinctive body of internal representations operating quasi-autonomously from the realm of direct sensorimotor engagement. Such representations are perceptual but are geared towards (and optimized for) the specific needs of reasoning and planning rather than those of fluent physical engagement. These representations are conditioned by a stream of inputs that do indeed originate at the sensors, but this stream proceeds in large part in parallel to the processing stream dedicated to the fluid control of online, fine-tuned, sensorimotor engagement, and is systematically insensitive to much of the lower-level detail.

These 'dual-stream' models appear to differ from strong sensorimotor models in at least two crucial respects. First, they depict visual experience as depending on a suite of representations optimized for reasoning and planning, whereas strong sensorimotor models depict visual experience as occurring when (possibly very fine-grained) sensorimotor knowledge is either simply *active*, or more plausibly when it is *put into contact with*, or *used for the purposes of*, reasoning and planning. Second, these models looks to be fully compatible with the idea (rejected outright by the strong sensorimotor model) that conscious visual experience might often (and perhaps always) depend on specific local aspects of internal representational activity rather than on whole-animal sensorimotor loops.

A major part of the empirical impetus for the dual-stream story comes from Milner and Goodale, who suggest (1995) that conscious visual awareness reflects information-processing activity in a specific visual processing stream geared towards enduring object properties, explicit recognition, and semantic recall. This

and making contact with it, as it were'. But such active probing surely characterizes the intelligent saccades of the driver's eyes even when the driver is attending to other matters and not consciously experiencing the details of the road. Alternatively, if active probing means something like 'probing in the context of attentive problem-solving', then we are back to the full-strength role for reason and planning assigned by O'Regan and Noë.

stream—the ventral stream—is also in charge whenever real-world objects are unavailable, and governs our attempts to mime actions on imagined or recalled objects. Actual object-based motor engagements, by contrast, are depicted as the province of a semi-autonomous processing stream—the dorsal stream—that guides fluent motor action in the here and now. Milner and Goodale thus contrast capacities of visually-guided action and capacities of conscious visual perception, suggesting that these come apart in a variety of unexpected and revealing ways.

In support of this hypothesis, Milner and Goodale invoke a rich body of data concerning patients with damage to areas in either the dorsal or ventral streams. The best known of these is the patient DF who suffers from ventral stream lesions and cannot visually identify objects or visually discriminate shapes. Nonetheless, she is able to pick up these very same objects—that she cannot visually identify—using fluent, well-oriented precision grips. Others, by contrast, suffer dorsal stream lesions and

> have little trouble seeing [i.e., identifying objects in a visual scene] but a lot of trouble reaching for objects they can see. It is as though they cannot use the spatial information inherent in any visual scene. (Gazzaniga 1998, p. 109)

Milner and Goodale also cite performance data from normal human subjects using experimental paradigms such as Aglioti et al.'s ingenious use of the Tichener Circles illusion. In the standard illusion, subjects misjudge the relative size of two circles each surrounded by a ring of larger or smaller circles. Aglioti et al. (1995) set up a physical version of the illusion using thin poker chips as the discs, and then asked subjects to 'pick up the target disc on the left if the two discs appeared equal in size and to pick up the one on the right if they appeared different in size' (Milner and Goodale 1995, p. 167). The surprising upshot was that even when subjects were unaware of—but clearly subject to—the illusion, their motor control systems produced a precisely fitted grip with a finger-thumb aperture perfectly suited to the *actual* (non-illusory) size of the disc. This aperture was not arrived at by touching and adjusting, but was instead the direct result of the visual input. Yet—to repeat—it reflected not the illusory disc size given in the subject's visual experience, but the actual size. In short:

> Grip size was determined entirely by the true size of the target disc [and] the very act by means of which subjects indicated their susceptibility to the visual illusion (that is, picking up one of two target circles) was itself uninfluenced by the illusion. (Milner and Goodale 1995, p. 168)

This was, indeed, a somewhat startling result, again suggesting[8] that the processing underlying visual awareness may be operating quite independently of that underlying the visual control of action.

Finally, Milner and Goodale invoke a number of computational conjectures concerning the inability of a single encoding to efficiently support both visual form recognition and visuomotor action. I shall not further rehearse these bodies of evidence here (see Clark 1999, 2001, Jacob and Jeannerod 2003). Notice, however, that it seems very likely that prediction learning will play a big role in the development and exercise of the kinds of skill that best characterize the (putatively non-conscious) dorsal stream. For smooth visuomotor action (such as reaching and grasping without jagged movements or unwanted oscillations) looks to require the ability to predict the next sensory input in advance of its actual (too late) arrival at the sensory peripheries, and this kind of prediction (see Clark and Grush 1999) certainly requires implicit knowledge of sensorimotor dependencies.

At this point it will be useful to locate the dual visual systems hypothesis in a wider framework. This framework depicts *conscious* visual perception as making available—at the personal level—forms of encoding and representation optimized for (or simply specialized for) their role in reasoning, choice and action-selection rather than for their role in actual sensorimotor engagement. Thus in the Titchener Circles experiment, the conscious visual representations would be specialized (just as Milner and Goodale suggest) to guide the choice of *which* disc to pick up, and the choice of what *kind* of grip to deploy (one apt for picking up and not for, say, throwing). The conscious illusion (of one circle's being larger than another) may then be best explained by the visual system's delivering a representation

8. There has been much discussion, pro and con, of this example in the recent literature. For a good overview (whose conclusion is that a weakened form of dual-stream model is probably correct) see Jacob and Jeannerod 2003. For further discussion of these complexities, see Clark 2001.

enhanced in the light of information about relative size, a trick that is effective for reasoning and choice in most ecologically realistic situations, but that would be damaging (resulting in a mass of failed or botched encounters) were it replicated by fine sensorimotor control systems.

This general model of the functional role of conscious awareness is found in (among others) the work of Koch (2004) who speaks of 'summaries' apt to aid frontal regions in the selection of one among a set of possible types of action or response. It is also suggested by Campbell's (2002) 'targets' view of consciousness, and (in the treatment closest to the present account) in Jacob and Jeannerod's (2003) delicately nuanced version of the dual visual streams view.[9]

Common to all these views is the image of conscious perceptual experience as reflecting the content of representations whose special cognitive role is to enable the deliberate selection of targets for action and of types of action, and to support a range of 'epistemic skills' such as sorting, sifting, comparing and the like (see Pettit 2003, Matthen 2005). Representations optimized for such purposes need not, and typically do not, reflect the full intricacies of our actual ongoing sensorimotor engagements with the world. Instead, they are geared, tweaked and nuanced so as to inform reason, selection, comparison and choice. They thus reflect only the very broad outlines of a space of possible targets and possible *kinds* of sensorimotor engagement. And though they must be sensitive to sensory input, they need not (indeed, must not) be sensitive to every nuance in the ongoing mass of sensory stimulation. The representations that ultimately determine visual experience are quite distinct, this alternative account insists, from those that support the sensorimotor loops by means of which we successfully engage the very world we perceive. They are, nonetheless, still distinctively visual, in so far as they represent the special kinds of information gathered (in normal agents) by (part of) the visual pathway: features such as rough spatial location, colour, shape, and so on. TVSS systems, on this account, aim to make the same kinds of information available by means of superficially different sets of signals, and will succeed to whatever extent this turns out to be

9. It also seems implicit in Matthen's (2005) account of the class of 'descriptive sensory systems'.

possible (which will in turn depend both on the nature and extent of neural plasticity and on the ability, of these alternate input devices, to make the same bodies of information available, and at roughly the same time-scale: for this take on TVSS, see Bach y Rita and Kercel 2003).

Noë (2004, p. 19) claims that these dual visual systems ideas are 'at best orthogonal to the basic claims of the enactive approach'. The reason given (see also op. cit., p. 11) is that the enactive approach makes no claims about what conscious visual perception is *for*, and hence remains neutral on the topic of vision-for-action versus vision-for-conscious-perception. More positively, O'Regan and Noë (2001, p. 969) claim that, with a few provisos, there is actually a good fit between the strong sensorimotor model and the dual visual systems ideas, since the requirement (for conscious experience) that sensorimotor knowledge be active in the service of reason and planning predicts the kinds of dissociation found in the literature.

I think it should be clear, however, that such direct attempts at reconciliation cannot succeed. For what is at issue is not simply the evidence of (substantial) dissociation but the best *functional and architectural explanation* of that evidence. And the best functional and architectural explanation, according to Milner, Goodale and others, is that conscious perceptual experience reflects the activation of representations that have less to do with the fine details of world-engaging sensorimotor loops, and more to do with the need to assign inputs to categories, types, and relative locations, so as better to sift, sort, select, identify, compare, recall, imagine and reason.

The contrast between the two views emerges in, for example, O'Regan and Noë's surprising description of DF as a case of 'partial awareness' in which 'she is unable to describe what she sees but is otherwise able to use it for the purpose of guiding action' (op. cit., p. 969). DF, recall, is able to use visually presented information for some purposes, for example, to post a letter into a slot, even though she claims not to visually experience the shape, colour or orientation of the slot. O'Regan and Noë depict this as a case of 'partial awareness', since visual information is still playing an action-guiding role in the overall organism–environment loop. But this surely conflates visual awareness with the use of visual information, precisely the knot that Milner and Goodale were trying to untie. For this reason Goodale (2001) rejects O'Regan and Noë's account of DF, pointing out that she 'shows almost perfect visuomotor control in the absence of any evidence that she actually "sees" the form of the object she is grasping'.

Here, I suspect, the enactive framework is trying to wag the empirical dog. For the enactive framework is, as we saw, pre-committed to linking the perceptual facts to facts about whole animal embedded, embodied activity. Perception, including conscious perception, is thus said to be 'a kind of skilful activity on the part of the animal as a whole' (Noe 2004, p. 2; and see Varela, Thompson and Rosch 1991).[10] But this pre-commitment works against taking truly seriously the evidence for deep dissociations between vision-for-action and vision-for-perception.

In contrast with this 'whole animal' view, dual stream models are open to the possibility that specific perceptual capacities and experiences depend upon (and can be brought about by) the activity of specific aspects of neural circuitry. In the case of conscious visual experience, such models embrace the idea that processing in the ventral stream plays a special role in the construction of conscious experience, and that there is serious functional decomposition (coupled with dense online integration; see Jacob and Jeannerod 2003) between systems for conscious experience and systems for fluent, fine-tuned visuomotor action.

Such models retain the important emphasis on skills rather than qualia. But they do so while recognizing the very large extent to which the human agent is a *fragmented bag of embodied skills*, only some of which are potentially relevant to the contents and character of perceptual experience. In particular, these will be skills geared rather directly towards reasoning and planning, such as abilities of sifting, sorting, classifying, selecting, choosing, re-identifying, imagining, recalling and comparing. This special focus opens up a significant buffer zone between the fine details of movement and of motion-dependent sensory input and the rather more specialized skill base that determines the contents

10. As an aside, this same broad commitment, to the constitutive role of whole organism activity, probably leads to other oddities, such as Noë's later suggestion that a concert pianist, in losing his arms, would thereby (instantly, as what appears to be a matter of conceptual necessity) lose his know-how since 'the knowledge was, precisely, arm-dependent' (Noë 2004, p. 121).

and character of perceptual experience. What counts (for perceptual experience) is then this suite of epistemic skills*, however they happen to be supported by cycles of low-level sensorimotor pick-up*. And there is, as far as I can see, no compelling reason to believe that *these* kinds of epistemic skills need to march in tight lockstep with a being's full sensorimotor profile. Indeed, they may depend on representational forms that are deliberately (that is to say, productively) insensitive to many fine details of bodily orientation and sensory stimulation. If this is correct, then the perceptual experience of differently embodied animals could in principle be identical (not merely similar) to our own.

I'd like to end by addressing (however inadequately) what is perhaps the very largest issue hereabouts, namely the role of actual world-engaging loops in the construction of perceptual experience. For it is here that the greatest care is needed if we are to preserve what is most important in the sensorimotor account. Thus consider, to take one final example, the claim (Noë 2004, p. 67) that 'experiential content is itself virtual'. The idea here, in keeping with the general emphasis on sensorimotor contact with the world, is that experience presents all the detail in a visual scene as present, but virtually so, rather like 'the way that a web site's content is present on your desktop' (op. cit., p. 50). In the latter case, it can seem just as if, to use Noë's own example, you have the entire contents of the online version of the *New York Times* encoded on your hard drive. But of course, this is not so. Rather, information is accessed from the distant site on a kind of just-in-time, need-to-know basis. Similarly, according to Noë, we perceptually experience the visual scene as rich in detail. But this experience, while not illusory (*pace* the Grand Illusion idea popularized by Dennett and others: see Noë 2004, pp. 50–67) is rooted not in the presence of a rich neurally encoded representation of all that detail, but in our skill-based access to the requisite detail as and when needed:

> The detail is present—the perceptual world is present—in the sense that we have a special kind of access to the detail, an access controlled by patterns of sensorimotor dependence with which we are familiar. (Noë 2004, p. 67)

This stress on access is correct and, I think, profoundly important.

But what *exactly* is the role of the actual sensorimotor loops by means of which such access is provided? How, that is to say, should we conceive the role of the *specific routines* by means of which we thus engage the world, retrieving more visual information as and when needed?

One radical possibility is that these specific sensorimotor loops are now part of the supervenience base for the present experience of richness. Another, only slightly less radical, possibility is that our implicit knowledge of these specific sensorimotor loops is part of the supervenience base for the present experience of richness. But still another possibility is that the present experience of richness is simply a present experience of the easy accessibility (of certain kinds of information) as and when needed, and that the specific world-engaging loops provide merely the contingent means to this end. The supervenience base for the perceptual experience of richness, on this model, would not include the routines that actually retrieve such information. Indeed, the very same experience of perceptual richness looks compatible with the running (behind the scenes, as it were) of a wide variety of quite different retrieval routines.

The deepest question raised by the strong sensorimotor model, is surely this: to what extent does the 'how' of information pick-up (the specific details of some sensorimotor retrieval routine) matter *for perceptual experience itself*? My own suspicion, which I have tried to make plausible in the present treatment, is that such details may be merely the contingent means by which a certain higher-level, overarching functional organization is achieved. The kind of functional story required will vary from case to case, but will typically be pitched at some remove from the full details of our active sensorimotor repertoire.

IV

The Heideggerian Theatre. In sum, none of the considerations adduced by the strong sensorimotor theorist seem to support the radical conclusion that qualitative experience is 'not caused by and realized in the brain' (Noë 2004, p. 226). What does seem to be true is that intentionally-driven cycles of world-engaging organism–environment activity are the typical means by which neural circuitry becomes recruited so as to be poised to control certain key kinds of action. It is this achieved poise, rather than anything more specific to the neural circuitry itself, that then does the real explanatory work. Thus the fans of skill-based approaches (and I count myself among them) believe that the right kinds of action-enabling poise (of, I want to say, some active inner state) quite literally constitute the presence of perceptual experience of various kinds. The considerations concerning ventral stream processing suggest, however, that the relevant kinds of action-enabling poise[11] are not tied to the fine detail of specific bodily equipment or motions. Instead, what matters is poise is for the control of a special class of more epistemically pregnant actions, such as those (highlighted by Dennett, Matthen and Pettit) of sifting, sorting, comparing and discriminating.

Suppose we then push the question: what is it about the active neural patterns that makes them yield or support the very experiences that they do? The best answer, just as the skill-based story insisted, looks to be: because they are such as to enable the kinds of activities they do. If that is so, there may simply be no answer to the bald (hard) question, 'Why does this pattern of neural pattern yield this experience?' Rather, the neural pattern, in bodily and worldly context, is such as to support a signature set of epistemic abilities, and it is these alone that explain (by constituting) the phenomenal content. The correct conclusion is that neural activity alone indeed suffices, as far as we can currently tell, to *support* any given perceptual experience, but that no explanation of the *link* between the neural activity and the experience can afford to ignore the shape of the space of enabled actions (of that special epistemic class) on pain of eventually identifying an unexplanatory disjunction instead of an explanatorily pregnant correlation.[12]

Strong sensorimotor models of perceptual experience do us a service by foregrounding embodied skills and eschewing appeals

11. It remains possible that more neglected elements of experience, ones other than those concerning the typical qualia suspects such as shape, colour, texture, and so on, may depend more directly on dorsal stream activity. Thus Matthen (2005, p. 301) argues that the 'feeling of presence' may depend on dorsal stream activity even if the other more descriptive elements do not.

12. For an argument of this kind, but without the stress on a special class of epistemic actions, see Hurley forthcoming.

to qualia as traditionally conceived. But they fail to do justice to the many firewalls, fragmentations and divisions of cognitive labour that characterize our engagements with the world our senses reveal. Strong sensorimotor models, despite (or perhaps because of) their emphasis on whole animal world-engaging activity, paper over this complex motley by casting everything prematurely in the single currency of known patterns of fine sensorimotor dependence. By trying to distil all that matters about human perceptual experience from the homogeneous mash of expectations concerning sensorimotor dependencies, such models are congenitally blind to the computationally potent *insensitivity* of key information-processing events to the full subtleties of embodied cycles of sensing and moving. In place of this common sensorimotor currency we need to consider a more complex picture that displays a cognitive economy replete with special-purpose streamings and with multiple, quasi-independent forms of internal (and external) representation and processing. To embrace such complexity is not to downplay the role of the body, but to reveal it aright. The body matters because it is the *common locus* of many (though by no means all) of these diverse epistemic currents and influences, and because it is their common, and often conflicted, object of proximal control. The body is (dare I say it?) the Heideggerian Theatre: the one place where it all comes together, or as together as it comes at all.

References

Aglioti, S., M. Goodale and J. F. X. DeSouza 1995: 'Size Contrast Illusions Deceive the Eye But Not the Hand'. *Curr. Biol.*, 5, pp. 679–85.

Bach y Rita, P. and S. W. Kercel 2003: 'Sensory Substitution and the Human–Machine Interface'. *Trends in Cognitive Sciences*, 7:12, pp. 541–6.

Ballard, D. H., M. M. Hayhoe, P. K. Pook and R. P. N. Rao 1997: 'Deictic Codes for the Embodiment of Cognition'. *Behavioral and Brain Sciences*, 20, 4, pp. 723–67.

Bridgeman, B., S. Lewis, G., Heit and M. Nagle 1979: 'Relation Between Cognitive and Motor-oriented Systems of Visual Position Perception'. *Journal of Experimental Psychology* (*Human Perception*), 5, pp. 692–700.

Campbell, J. 2002: *Reference and Consciousness*. Oxford: Oxford University Press.

Churchland, P., V. Ramachandran and T. Sejnowski 1994: 'A Critique of Pure Vision'. In C. Koch and J. Davis (eds), *Large-Scale Neuronal Theories of the Brain*. Cambridge, MA: MIT Press, pp. 23–61.

Clark, A. 1997: *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA: MIT Press.

——— 1999: 'Visual Awareness and Visuomotor Action'. *Journal Of Consciousness Studies*, 6, 11–12, pp. 1–18.

——— 2000: 'A Case Where Access Implies Qualia?' *Analysis*, 60:265, pp. 30–8.
——— 2001: 'Visual Experience and Motor Action: Are The Bonds Too Tight?' *Philosophical Review*, 110:4, pp. 495–519.
——— and R. Grush 1999: 'Towards a Cognitive Robotics'. *Adaptive Behavior*, 7:1, pp. 5–16.
——— and P. Mandik 2002: 'Selective Representing and World-Making'. *Minds and Machines*, 12, pp. 383–95.
——— and J. Toribio 2001: 'Sensorimotor Chauvinism?: Commentary on O'Regan, J. K., and Noë, A., "A sensorimotor approach to vision and visual consciousness" '. *Behavioral and Brain Sciences*, 24:5, pp. 979–80.
Dennett, D. 1991: *Consciousness Explained*. Boston, MA: Little, Brown.
Elman, J. 1995: 'Language as a Dynamical System'. In R. Port and T. van Gelder (eds), *Mind as Motion*. Cambridge, MA: MIT Press.
Gazzaniga, M. 1998: *The Mind's Past*. Berkeley, CA: University of California Press.
Gibson, J. J. 1979: *The Ecological Approach to Visual Perception*. New York: Houghton-Mifflin.
Goodale, M. 2001: 'Real Action in a Virtual World: Commentary on O'Regan and Noë'. *Behavioral and Brain Sciences*, 24:5, pp. 984–5.
Heidegger, M. [1927] 1961: *Being and Time*. London: Harper and Row.
Hurley, S. 1998: *Consciousness in Action*. Cambridge, MA: Harvard University Press.
——— forthcoming: 'The Varieties of Externalism'. In M. Aydede and P. Robbins (eds), *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press.
Husserl, E. 1907: *Thing and Space*. Trans. R. Rojcewicz. Boston, MA: Kluwer.
Hutchins, E. 1995: *Cognition in the Wild*. Cambridge, MA: MIT Press.
Jacob, P. and M. Jeannerod 2003: *Ways of Seeing: The Scope and Limits of Visual Cognition*. Oxford: Oxford University Press.
Koch, C. 2004: *The Quest for Consciousness*. New York: Roberts and Co.
MacKay, D. 1967: 'Ways of Looking at Perception'. In W. Wathen-Dunn (ed.), *Models for the Perception of Speech and Visual Form*. Cambridge, MA: MIT Press, pp. 25–43.
Matthen, M. 2005: *Seeing, Doing and Knowing*. Oxford: Oxford University Press.
Meijer, P. B. L. 1992: 'An Experimental System for Auditory Image Representations'. *IEEE Transactions on Biomedical Engineering*, 39:2, pp. 112–21.
Merleau-Ponty, M. [1945] 1962: *The Phenomenology of Perception*. Trans. Colin Smith. London: Routledge.
Milner, A. and M. Goodale 1995: *The Visual Brain in Action*. Oxford: Oxford University Press.
Noë, A. 2004: *Action in Perception*. Cambridge, MA: MIT Press.
O'Regan, J. K. 1992: 'Solving the "Real" Mysteries of Visual Perception: The World as an Outside Memory'. *Canadian Journal of Psychology*, 46:3, pp. 461–88.
——— and A. Noë 2001: 'A Sensorimotor Approach to Vision and Visual Consciousness'. *Behavioral and Brain Sciences*, 24:5, pp. 939–73.
Pettit, P. 2003: 'Looks as Powers'. *Philosophical Issues*, 13, pp. 221–52.
Ryle, G. [1949] 1990: *The Concept of Mind*. London: Penguin.
Sur, M., A. Angelucci and J. Sharma 1999: 'Rewiring Cortex: The Role of Patterned Activity in Development and Plasticity of Neocortical Circuits'. *Journal of Neurobiology*, 41:1, pp. 33–43.
Thelen, E. and L. Smith 1994: *A Dynamic Systems Approach To The Development Of Cognition And Action*. Cambridge, MA: MIT Press.
Varela, F., E. Thompson and E. Rosch 1991: *The Embodied Mind*. Cambridge, MA: MIT Press.