WHERE EXPERIENCES ARE: DUALIST, PHYSICALIST, ENACTIVE AND REFLEXIVE ACCOUNTS OF PHENOMENAL CONSCIOUSNESS

Max Velmans, Department of Psychology, Goldsmiths, University of London, New Cross, London SE14 6NW; email m.velmans@gold.ac.uk
web address http://www.goldsmiths.ac.uk/departments/psychology/staff/velmans.html

Phenomenology and the Cognitive Sciences (in press)

Abstract.
Dualists believe that experiences have neither location nor extension, while reductive and 'non-reductive' physicalists (biological naturalists) believe that experiences are really in the brain, producing an apparent impasse in current theories of mind. Enactive and reflexive models of perception try to resolve this impasse with a form of "externalism" that challenges the assumption that experiences must either be nowhere or in the brain. However, they are externalist in very different ways. Insofar as they locate experiences anywhere, enactive models locate conscious phenomenology in the dynamic interaction of organisms with the external world, and in some versions, they reduce conscious phenomenology to such interactions, in the hope that this will resolve the hard problem of consciousness. The reflexive model accepts that experiences of the world result from dynamic organism-environment interactions, but argues that such interactions are preconscious. While the resulting phenomenal world is a consequence of such interactions, it cannot be reduced to them. The reflexive model is externalist in its claim that this external phenomenal world, which we normally think of as the "physical world," is literally outside the brain. Furthermore, there are no added conscious experiences of the external world inside the brain. In the present paper I present the case for the enactive and reflexive alternatives to more classical views and evaluate their consequences. I argue that, in closing the gap between the phenomenal world and what we normally think of as the physical world, the reflexive model resolves one facet of the hard problem of consciousness. Conversely, while enactive models have useful things to say about percept formation and representation, they fail to address the hard problem of consciousness.

Keywords: dualism, physicalism, enactive, reflexive, phenomenology, consciousness, externalism, internalism, reductionism, consciousness, mind, brain, world, perception, Noe, Thomson, Velmans, O'Regan, Myin, projection, space, phenomenal world, Lehar

The problems of consciousness have been extensively discussed in the scientific and philosophical literature. What is it? What is its function? And how are we to understand the causal relations between consciousness and brain? In Velmans (2000, 2003a) I have argued that these problems are partly empirical and partly conceptual. Why is the nature of consciousness partly a conceptual problem? We already have innumerable examples of conscious experience in our everyday lives and still puzzle over its nature, so merely having some more experiences, and thinking about them in our habitual ways, won't advance our understanding of its nature. Dualists and reductionists, for example, have

persistently differing opinions about the nature of consciousness in spite of the commonalities of their intersubjectively shared experiences.

The dualist view

The dualist view, which many people intuitively adopt, is shown in schematic form in Figure 1 below.

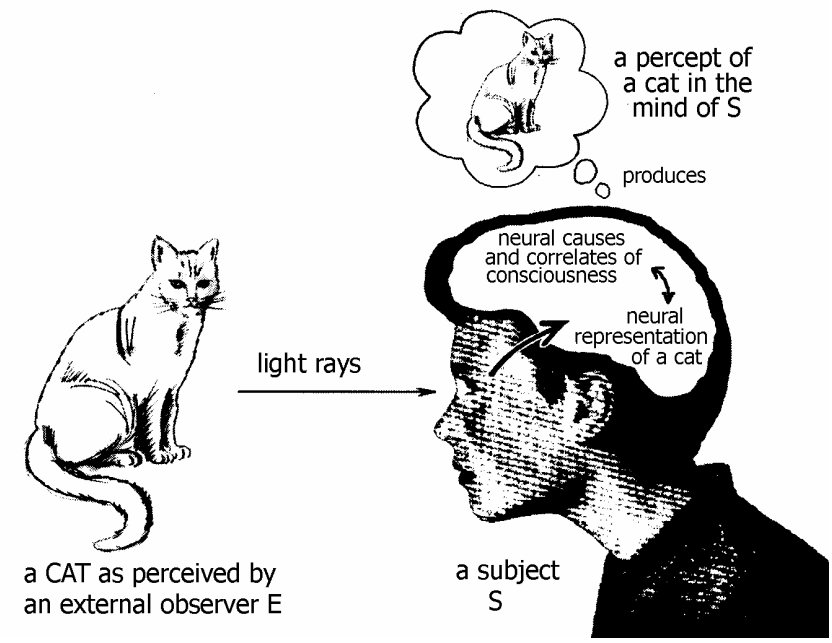

Figure 1. A dualist model of perception (adapted from Velmans, 2000)

This assumes perception to involve a simple, linear, causal sequence. Viewed from the perspective of an external observer E, light rays travelling from the physical object (the cat as-perceived by E) stimulate the subject's eye, activating her optic nerve, occipital lobes, and associated regions of her brain. Neural conditions sufficient for consciousness are formed, and result in a conscious experience (of a cat) in the subject's mind. This model of visual perception is, of course, highly oversimplified, but for now we are not interested in the details. We are interested only in where external physical objects, brains and experiences are *placed.*

It will be clear that there are two fundamental "splits" in this model. Firstly, the contents of consciousness are clearly separated from the material world (the conscious, perceptual "stuff" in the upper part of the diagram is separated from the material brain and the physical cat in the lower part of the diagram). This conforms to Descartes' view that stuff of consciousness (*res cogitans*, a substance that thinks) is very different to the stuff of which the material world is made (*res extensa*, a substance that has extension and location in space). Secondly, the perceiving *subject* is clearly separated from the perceived *object* (the

subject and her experiences are on the right of the diagram and the perceived object is on the left of the diagram).

This dualist model of perception supports a dualist view of the universe in which the universe is split into two realms, the material realm and the mental realm (the latter including consciousness, mind, soul and spirit). In interactionist forms of dualism these two realms interface and causally interact somewhere in the human brain.

The reductionist view

The problems of assimilating such dualism into a scientific worldview are serious (cf Velmans, 2000 ch2). Consequently, it is not surprising that 20th Century philosophy and science tried to naturalise dualism by arguing or attempting to show that conscious experiences are nothing more than states or functions of the brain. A reductionist model of visual perception is shown in Figure 2.

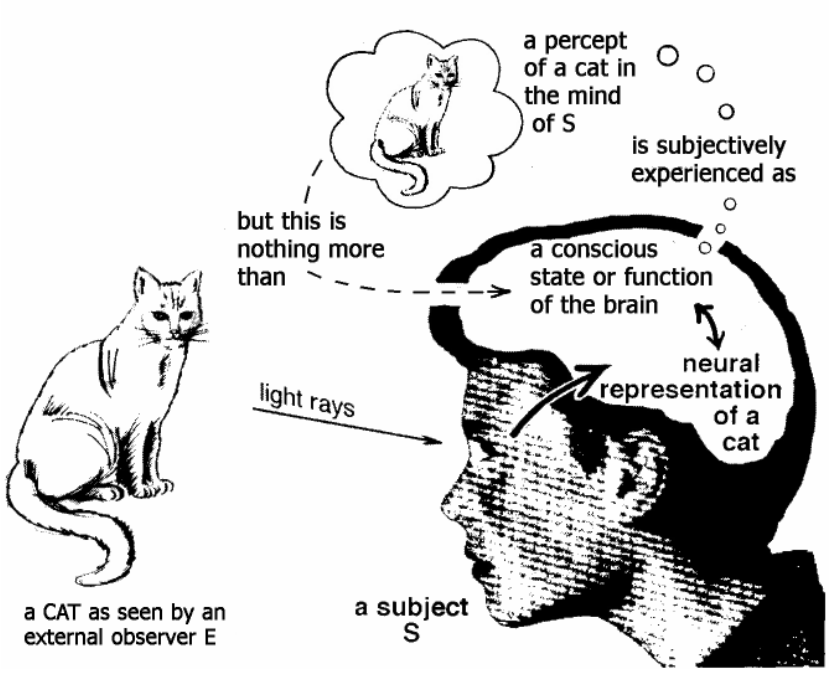

Figure 2. A reductionist model of perception (adapted from Velmans, 2000)

The causal sequence in Figure 2 is the same as in Figure 1, with one added step. While reductionists generally accept that the subject's experience of a cat *seems* to be insubstantial and "in the mind", they argue that it is *really* a state or function of the brain. In short, the reductionist model in Figure 2 tries to resolve the conscious experience—physical world split by eliminating conscious experience or reducing it to something physical that E (the external observer) can in principle observe and measure. But reductionism *retains* the split (implicit in dualism) between the observer and the observed. The perceived object (on the

left side of the diagram) remains quite separate from the conscious experience *of* the object (on the right side of the diagram).

This supports a reductionist view of a universe entirely composed of physical material, of which conscious experiences are a tiny part (the bits of human brain identified with those experiences).

Note that in spite of their disagreement about the ontology of conscious experiences, dualists and reductionists largely agree about how conscious experiences *relate* to the brain and physical world. In visual perception, for example, they would agree that physical input stimuli innervate the optic nerve and visual system, forming preconscious representations of that input in the brain. If that input is attended to, and the necessary and sufficient conditions for consciousness are met, a conscious experience will result along with its neural correlates in the brain. This agreement further highlights the conceptual nature of the dualist versus reductionist debate. If it is agreed that visual experiences have neural causal antecedents as well as neural correlates, the empirical *discovery* of those causes and correlates won't settle their dispute about whether the experiences are *nothing more than* their causes and/or correlates.

Note too that dualists and reductionists largely agree about where the external physical world, the brain and conscious experiences are *placed.* In spite of their dispute about *what* experiences are, they agree (roughly) about *where* they are. Reductionists, for example, take it for granted that experiences are really brain states or functions, so they must be *in the brain.* Although dualists take experiences to be immaterial (and, strictly speaking, without location or extension) they again take it for granted that these must interface and interact with the physical world somewhere in the brain. In short, the brain is as close to experiences as one can get—and if experiences are in the brain, they cannot be located in, or part of, the external physical world. One could describe this view as *phenomenological internalism*.

Enactive and *reflexive* approaches to consciousness challenge the grounds of the dualist versus reductionist debate by challenging some of these seemingly innocuous assumptions that dualists and reductionists *share*, namely a) what conscious experiences *seem to be like*, and b) *where* conscious experiences are placed in relation to the brain and the physical world.

Enactive accounts of consciousness

There are a family of theories that might broadly be described as enactive, characterised by an emphasis on perception being a sensory-motor skill involving ongoing interaction with the external world, rather than being dependent on an "inner representation" of that world. These theories can be considered in terms of what they say about the way that perceptual processing works, and in terms of what they say about the ontology of conscious experiences (for example, about the nature of "qualia"). While questions about perceptual functioning and about conscious phenomenology are, in principle, separable, a number of enactive theorists claim them to be connected: according to them, if one

understands perceptual functioning in an enactive way as mastery of a set of sensory-motor skills, one can also understand the nature of conscious experience including its "qualia" in this way—thereby (hopefully) resolving one of the hard problems of consciousness (see, e.g. O'Regan and Noe, 2001; O'Regan, Myin and Noe, 2005).

At one extreme, theorists sympathetic to the enactive approach argue that a better understanding of how, say, visual perception works, resolves the hard problem of qualia by showing our beliefs about our visual experiences to be entirely false. For example, Dennett (2002) points out that we commonly think that we experience the visual world in fine detail and colour from the centre of the visual field to the periphery. However, science demonstrates that this cannot be so, having discovered that peripheral vision has poor acuity and does not code for colour. According to Dennett, if we can be wrong about colour extending to our visual periphery, we can be wrong about everything to do with our phenomenology, for example, that we experience colour qualia at all. In the same vein, Blackmore (2002) claims "there is no point trying to explain the differences between things that are in consciousness and those that are not because there is no such difference. And it is a waste of time trying to explain the contents of the stream of consciousness because the stream of consciousness does not exist." (p28)

Other enactive theorists however, would regard such views as unjustifiably extreme. While *some* people might have false beliefs about some aspects of their experience that they have not bothered to think about closely, it does not follow that *all* of our beliefs about our experiences are or need be wrong—and certainly does not warrant the claim that conscious qualia do not exist! A moment's close attention to the qualia of our visual field, for example, is all that is needed to confirm the description of it given by science. Once one attends to it, it is perfectly obvious that we can discern fine detail at our point of focus but not at the periphery of vision, and that although colour and detail at the periphery are at best fuzzy, both colour and detail at the focus of attention are clear. In short, lack of detail and colour of qualia at the periphery of vision has no bearing on the existence of detail and colour of qualia at the focus of vision, and therefore no bearing on the existence of qualia as such.[1]

Recent findings on inattentional and change blindness are, however, more challenging to our everyday beliefs. Studies of inattentional blindness such as Simons & Chabris (1999), for example, suggest that we do not see what we do not attend to *even when we are directing our gaze at it*. Equally surprising, studies of change blindness such as Simons & Levin (1998) demonstrate that we do not notice *major changes* in what we are gazing at unless fast transitions capture our attention, or we happen to be focusing our attention on the precise features that change. Taken together, such findings provide persuasive demonstrations that what we notice about the perceived world is less complete and detailed than we usually think. The findings also challenge a commonly held view within psychology about how perception works, namely that we have a detailed, and complete inner representation of the external world built up over successive eye saccades out of the degraded information arriving at the retinas. If such a complete representation were updated moment-by-moment, then we should notice changes in the visual field by

comparing current input with complete records of the world developed from prior input—but we don't.

The alternative, *enactive* view suggests that we perceive perhaps 5 to 6 features of the world at any given moment (wherever we gaze) but we are free to pick up any other features, as we need them, by exploring the world (e.g. with eye movements). The reason that we think that the visual world is rich in detail and colour is because the world itself does have this detail and colour, and we see this wherever we look. We do not need to build up a complete, detailed inner representation of the world because the world itself stores all the relevant information.

If true this would be a genuine advance in our understanding of how perception works (that we pick up just 5 or 6 visual features at each fixation) and about the nature of consequent inner representations of the world (that they are limited to the features that are picked up and are, therefore, not complete). The dynamic interaction between internal information and external information (picked up on a need to know basis) also suggests that internal information may sometimes be formatted in a way that is suited to such ongoing activities, for example as a set of procedures for action, rather than being iconic or propositional. The idea that inner representations are at least in part procedural rather than iconic or propositional is a recurring theme in cognitive science (see for example the procedural semantics developed in considerable depth by Miller & Johnson-Laird, 1976). However, the inattentional and change blindness data does not suggest that there are *no inner representations at all* (an extreme form of "externalism" sometimes associated with the enactive view, see, e.g. Noe & Thomson, 2004a). Nor does any of this data suggest that there are no "qualia." On the contrary, there are qualia (associated with the 5 to 6 features we pick up) wherever we look.

A group of enactive theorists nevertheless claim that understanding perception in this way allows an *explanation* of the qualia of consciousness in ways that are not open to the more traditional view that these are, in some mysterious fashion, generated by neural activity in the brain. For example, O'Regan, Myin and Noe (in press?) ask, "What is it exactly about phenomenal consciousness which makes it seem inaccessible to normal scientific inquiry? What is so special about "feel"?" Their reply is that "Feel is…not "generated" by a neural mechanism at all, rather, it is exercising what the neural mechanism *allows the organism to do*." The feel of driving a Porsche for example does not reside in any given moment, but rather in the fact that you are currently engaged in exercising the Porsche driving skill. And, "If the feel of Porsche driving is constituted by exercising a skill, perhaps the feel of red, the sound of a bell, the smell of a rose also correspond to skills being exercised." Readers familiar with the consciousness studies literature will recognise that this reductive identification of conscious "feel" with the exercising of a sensory-motor skill is a variant of functionalism, although it locates the relevant functioning in the skilful interaction of organisms with the surrounding world rather than in causal relationships that are exclusively located within the brain. Consequently it may be seen as a variant of the position shown in Figure 2, although it is difficult to represent the exercising of sensory-motor skills in a simple, schematic diagram (I will leave this to the reader as an exercise).

The notion that sensory-motor interaction with the world may affect aspects of visual perception that have to do with apparent size, shape, location and orientation of objects in space again has classical antecedents in psychology, dating back to experiments with inverted retinal images in 1897 by G.M. Stratton, and work with a variety of distorting spectacles by Ivo Kohler and others in the 1960s (see Velmans, 2000, ch7). It was also demonstrated, for example, by the way that the distorted appearance of an Ames room[2] only changes once one interacts with it, e.g. by walking around it, or poking the sides of the room with a stick (the appearance is not changed if one is simply told its true dimensions). That said, many would doubt that conscious qualia that cannot be explored with motor movements (such the appearance of red, the sound of a bell or the smell of a rose) are similarly dependent on sensory-motor skills. The nose, for example, has few motor options: one can wrinkle it and point it in different directions, but to the best of our knowledge it is its ability to fit appropriately shaped vapour molecules into appropriately shaped receptor sites, not nasal sensory-motor skill, that enables it to discriminate the smell of a rose from the smell of ripe cheese. However, I will leave the fuller analysis of this more dubious, theoretically overextended aspect of the enactive view to other commentators in this double issue. I want to focus on something more fundamental—on whether sensory motor skill gets one any closer to explaining the hard problem of conscious "feel."

Why should driving a Porsche or any other skill *feel like anything at all*? I am not denying that functioning of different kinds in humans often feels like something for humans. However human functioning can often be dissociated from its normal feel. For example, once they are well learnt, consciously performed skills can often be performed unconsciously,[3] so it does not follow that skilful functioning itself *explains* the accompanying feel.[4]

If it is a *contingent*, not a *necessary* fact that certain kinds of functioning in humans have certain kinds of feel, then switching one's emphasis away from neural mechanisms as such, to "what neural systems allow an organism to do" gets one no closer to understanding why that enabling of skill should have a feel at all. Piloting a 747 no doubt, feels like something, to *a human pilot* and the way that it feels is likely to have something to do with human biology. But why should it feel the same way to an electronic autopilot that replaces the skills exercised by a human being? Or why should it feel like anything to be the control system of a guided missile system? Anyone versed in the construction of electronic control systems knows that if one builds a system in the right way, it will function just as it is intended to do, *whether it feels like anything to be that system or not.* If so, functioning in an electronic (or any other) system is logically tangential to whether it is like anything to be that system, leaving the hard problem of why it happens to feel a certain way in humans untouched[5].

In short, if one *helps oneself* to the feelings that accompany certain sensory motor acts it might be possible, or not, to extend some of those feelings to aspects of conscious experience that are not normally associated with skilful acts, thereby persuading us that these are qualitatively different to how we normally think them to be, as some enactive theorists propose. But this tells us nothing about why skilful acts themselves should feel

like anything at all, and consequently fails to address the hard problem of conscious "qualia."

To sum up, the enactive view differs from standard physicalism and functionalism in that it replaces neural representations in the subject's brain with sensory-motor interactions between subject and external world that cross the subject—object divide, and it is in this sense "externalist". However, like physicalists and functionalists, some enactive theorists try to resolve the problem of "qualia" and with it, the conscious experience/physical world split, by reducing conscious experiences to something that E (the external observer) can in principle observe and measure (to the exercise of sensory-motor skills). To the extent that they try to reduce how things appear from a subject's first-person perspective to how things appear from E's third-person perspective they are reductionist. I have given some initial reasons to doubt the viability of this enactive alternative to more standard forms of reductionism above—and reductionism has many, additional problems that I do not have space to elaborate on here (cf. Velmans, 2000 chs.3, 4 & 5). Instead, I want to introduce a more radical proposal: that the dualist and reductionist models of perception shown in Figures 1 and 2 above, should be replaced by the reflexive model of perception shown in Figure 3.

The Reflexive Model of Perception

Like the enactive view, the reflexive model of perception proposes a different analysis of how conscious experience relates to the brain and surrounding world (Velmans, 1990, 2000 ch6). It also accepts that certain forms of perception arise from a dynamic interaction of observer with observed and that at least some aspects of this interaction have a sensory-motor component. However, unlike some supporters of the enactive view, it assumes that such sensory-motor interactions with the world are normally *preconscious*: these interactions may form part of the causal antecedents to a given experience, but antecedent causes are not the same as their consequent effects. Consequently, these sensory-motor interactions cannot be ontologically identical to the resulting experiences, and cannot (in this ontological sense) explain their nature (cf. Velmans, 1998, 2000 ch3).

Nor does the reflexive model argue against the existence or importance of internal neural representations, although it remains open about the nature of such representations, which may be iconic, procedural or in some other format. Whatever the format, internal representations are required to support memory, imagery, dreams, and hallucinations (where there is no external stimulus with which to engage in sensory-motor exploration) and once activated, these representations can also be sufficient, proximal causes for very detailed experiences.[6] In short, the reflexive model adopts a largely conventional approach to the *causes* of perception, while accepting that our knowledge of how perceptual processing works is, at present, partial, and needs to be informed in its details by data about change and inattentional blindness (along the lines suggested above) and by whatever other findings emerge. However, even a full understanding of preconscious

causes and *correlates*, described from a third-person perspective, won't tell us all that we need to know about the consequent *effects*, the qualia of consciousness. To know about these, we have to ask the subject, and it is only when we have such first-person data that we can construct a complete model of perception. One model of perception that combines (the third-person) information available to an experimenter with an accurate description of what the subject experiences is shown in Figure 3.

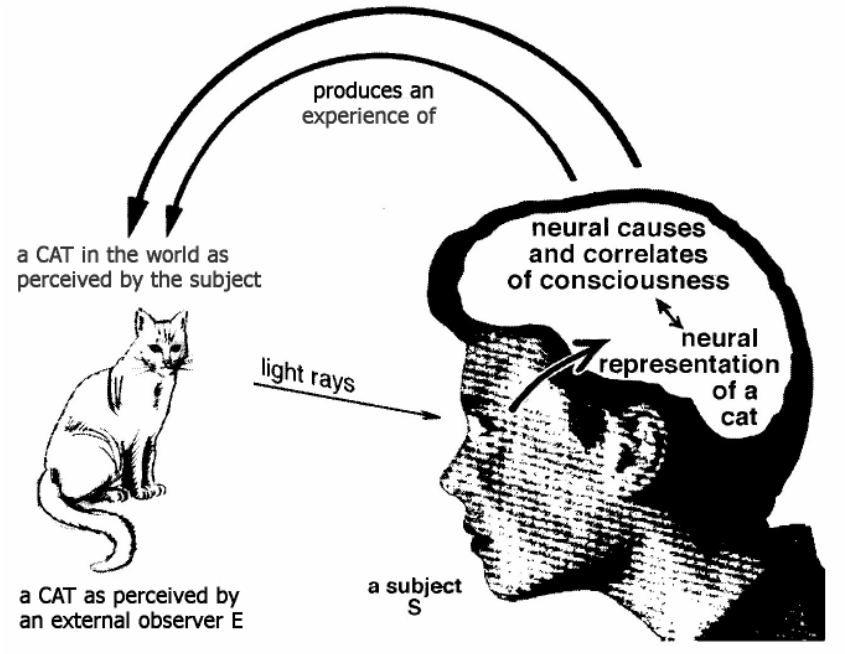

Figure 3. A reflexive model of perception
(adapted from Velmans, 2000)

In most respects Figure 3 is the same as Figures 1 and 2. As before, there is a cat in the world (perceived by E) that is the initiating stimulus for what S observes, and the proximal neural causes and correlates of what S experiences are, as before, located in S's brain.[7] The only difference relates to the ontology and location of S's experience. According to dualists, S's experience of a cat consists of "stuff that thinks" that is located "nowhere"; according to reductionists, S's experience of a cat is a state or function of the brain that is located in her brain; according to the reflexive model, both of the former models are theoretically rather than empirically driven with the consequence that they systematically misdescribe what S actually experiences. If you place a cat in front of S and ask her to describe what she experiences, she should tell you that she sees a cat in front of her in the world. This phenomenal cat literally *is* what she experiences, located where it seems to be—and she has no *additional* experience *of* a cat either "nowhere" or

"in her brain." According the reflexive model, this added experience is a myth. Applying Occam's razor gets rid of it.

In short, the reflexive model's *externalism* applies to the *phenomenology* of some experiences. Unlike the externalism of enactive theory, which applies to the antecedent causes or vehicles of given experiences, the central claim of the reflexive model is that insofar as experiences are anywhere, *they are roughly where they seem to be*. For example, a pain in the foot really is in the foot, and this perceived print on this page really is out here on this page. Nor is a pain in the foot accompanied by some *additional* experience *of* pain in the brain, or is this perceived print accompanied by some additional experience *of* print in the brain. In terms of phenomenology, this perceived print, and my experience *of* this print are *one and the same.*[8]

It should be easy to grasp the essence of this. The external objects that we experience seem to be out there in the world, not in our head or brain—and classical "mental" sensations such as itches and pains seem to be clearly located on the surface of our skin. But this immediately presents us with the problem of *perceptual projection*: given that the proximal neural causes and correlates of what we experience *are* in the head or brain, how can we explain the fact that various sensations and experiences seem to be beyond the brain?

Perceptual projection

It is important to be clear about what is meant by "perceptual projection" in order to convey its role in the reflexive model. Crucially, perceptual projection refers to an *empirically observable effect*, for example, to the fact that this print seems to be out here on this page and not in your brain. In short, perceptual projection is an effect that requires explanation; perceptual projection is not itself an explanation. We know that preconscious processes within the brain produce consciously experienced events, which may be subjectively located and extended in the phenomenal space beyond the brain, but we don't really know how this is done. We also know that this effect is subjective, psychological and viewable only from a first-person perspective. Nothing *physical* is projected from the brain. Although we don't have a full understanding of how perceptual projection works, there is a large experimental literature about the information that is used by the brain to model distance and location. There are also many ways to demonstrate perceptual projection in action, for example in hallucinations, phantom limbs, stereoscopic pictures, holograms, and virtual realities. I have discussed this literature elsewhere, along with some potentially useful models (holography and virtual reality) in Velmans (1990, 2000); but for our present purposes we do not need to examine the details. We simply need to note that the evidence for perceptual projection is all around us. In spite of the fact that the proximal neural causes and correlates of conscious experiences are inside our brains, our experienced phenomenal bodies and worlds appear to be outside our brains.

How phenomenal space relates to real space

No one doubts that physical bodies can have real extension and location in space. Dualists and reductionists nevertheless find it hard to accept that experiences can have a real, as opposed to a 'seeming' extension and location. They do not doubt, for example, that a foot has a real extension and location in space, but, for them, a pain in the foot can't really be in the foot, as they are committed to the view that it is either nowhere or in the brain. In sum, location in phenomenal space is not location in real space.

According to the reflexive model however, this ignores the fact that, in everyday life, we take the phenomenal world to *be* the physical world. It also ignores the pivotal role of phenomenal space in forming our very understanding of space, and with it, our understanding of location and extension in measured or "real" space.
What we normally think of as the "physical foot" for example is actually the *phenomenal foot* (the foot as seen, felt and so on). That does not stop us from pointing to it, measuring its location and extension and so on. If so, at least some phenomenal objects can be measured. While a pain in the foot might not be measurable with the same precision, few would doubt that we could specify its rough location and extension (and differentiate it for example from a pain in the back).

What we normally think of as "space" also refers, at least in the initial instance, to the phenomenal space that we experience through which we appear to move. Our intuitive understanding of spatial location and extension, for example, derives in the first instance from the way objects and events appear to be arranged relative to each other in phenomenal space (closer, further, behind, in front, left, right, bigger, smaller and so on). We are also accustomed to making size and distance estimates based on such appearances. This print for example appears to be out here in front of my face, and THIS PRINT appears to be bigger than this print. However, we recognise that these ordinal judgments are only rough and ready ones, so when we wish to establish "real" location, distance, size or some other spatial attribute, we usually resort to some form of *measurement* that quantifies the dimensions of interest using an arbitrary but agreed metric (feet, metres etc), relative to some agreed frame of reference (for example a Cartesian frame of reference with an agreed zero point from which measurement begins). The correspondence, or lack of correspondence, between phenomenal space and measured space is assessed in the same way (e.g. by comparing distance judgments with distance measurements) in psychology experiments. For example, I can estimate the distance of this phenomenal print from my nose, but I can also place one end of a measuring tape on the tip of my nose (point zero) and the other end on this print to determine its real distance.[9]

Such comparisons allow one to give a broad specification of how well phenomenal space corresponds to or maps onto measured space. According to the reflexive model, phenomenal space provides a natural representation, shaped by evolution, of the distance and location of objects viewed from the perspective of the embodied observer, which models real distance and location quite well at close distances, where accuracy is important for effective interaction with the world. My estimate that this page is about 0.5 metres from my nose, for example, is not far off. However, phenomenal appearances and our consequent distance judgments quickly lose accuracy as distances increase. For

example, the dome of the night sky provides the outer boundary of the phenomenal world, but gives a completely misleading representation of distances in stellar space.[10]

Note that, although we can use measuring instruments to correct unaided judgments of apparent distance, size and so on, measuring tapes and related instruments themselves appear to us as phenomenal objects, and *measurement operations appear to us as operations that we are carrying out on phenomenal objects in phenomenal space*. In short, even our understanding of "real" or measured location is underpinned by our experience of phenomenal location. And crucially, whether I make distance judgments about this perceived print and judge it to be around 0.5 metres in front of my face, or measure it to find that it is only 0.42 metres, *does not alter the phenomenon that I am judging or measuring.* The distance of the print that I am judging or measuring is the distance of this perceived print out here on this page, and not the distance of some other 'experience of print' in my brain.

Why this matters

These observations about the spatially extended nature of the experienced phenomenal world fit in with common sense and common experience and they will come as no surprise to those versed in European phenomenology. They also have many theoretical antecedents, for example in the work of Berkeley, Kant, and Whitehead, the neutral monism of James, Mach, and Russell, and the scientific writings of Köhler and Pribram. However, the view that experienced phenomena which seem to be outside the brain really *are* outside the brain is fiercely resisted by both dualists and reductionists, who are committed to the view that experiences must either be nowhere or in the brain, and therefore *separate* from what we normally think of as "the external world". Enactive theorists are similarly eager to distance themselves from a claim about conscious phenomenology that is supposedly so 'unscientific.' Noe & Thomson (2004b) for example make it clear that they do not defend externalism about experience (as they have different views on this matter). And they wish to defend themselves against the charge made by Jack & Prinz (2004) that "In the hands of Noe & Thomson, externalism becomes an eccentric doctrine that locates consciousness outside 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" (p55) In their defence, Noe & Thomson (2004b) stress that "It is *not* our view that consciousness is outside the head, but rather that some of the causal substrates of consciousness might be." (p93)

Yet, according to the reflexive model, it is precisely in the confused, unempirical, and doctrinal nature of some philosophical and so-called scientific thinking on this issue that a major source of the hard problem of consciousness is to be found. To understand how conscious experience relates to the brain and physical world, one must first describe the phenomenology of that experience *accurately*. If conscious phenomenology is systematically misdescribed, its relation to the brain and physical world cannot be understood. The empirical fact of the matter appears to be that preconscious processing in

the embodied brain interacting with the world results in the three-dimensional, external phenomenal world that we experience. In everyday life, it is precisely this 3D phenomenal world that we see, hear, touch, taste and smell around our bodies that we *think of as the physical world*, although we recognise that this experienced physical world only models in a rough and ready way the subtler world described by modern physics (in quantum mechanics, relativity theory, etc.). If so, there never was an explanatory gap between what we normally think of as the physical world, and conscious experience. This phenomenal physical world is *part of* conscious experience, not *apart from* it.

It should be apparent that this observation, if true, would alter the nature of the "hard problem" of consciousness, although, in isolation, it can be no more than a first step on the way to a theory. There is more than one thing to understand, for example the relation of conscious qualia to their neural correlates, the relation of the phenomenal physical world to the world described by modern physics, the causal efficacy and function of consciousness, and so on. I do not have space to present a more detailed theory here, although I have done so elsewhere (see, for example, Velmans, 1990, 2000, chs. 6 to 12, 2003). I will, however, try to make it clear why this first step is crucial.

Is the brain in the world or the world in the brain?

Readers familiar with the problem of conscious 'location' will recognise that the force of my suggestion that some experiences have both a spatial location and extension outside the head hangs on whether the *appearance-reality* distinction can be applied to conscious phenomenology. Are experiences *really* where they *seem* to be or not?

Although various thinkers have noticed the apparent spatial location and extension of some experiences, and have tried to fit this into a general theory of mind (see above), few workers in modern consciousness studies have noted the potential consequences of this for an understanding of consciousness. Of those that have, some have tried to dismiss the significance of spatially extended phenomenology with the argument that, if the neural causes of experience are in the brain, the experiences themselves must be there too. However, this presupposes the truth of a local model of causation that has long been abandoned by physics (which accepts that electricity inside a wire can cause a magnetic field outside the wire, that planets exert a gravitational pull on each other at great distances, that there are non-local effects in quantum mechanics, and so on).

Of more interest are a number of thinkers who take the apparent, spatially extended nature of much of experience very seriously, but nevertheless argue that such experiences are really brain states that are by definition in the brain. As it turns out, their attempt to assimilate 3D phenomenology into a form of "biological naturalism" is highly instructive.

In the modern era, John Searle was one of the first to address this problem. As he noted,

"Common sense tells us that our pains are located in physical space within our bodies, that for example, a pain in the foot is literally in the physical space of the foot. But we now know that is false. The brain forms a body image, and pains like all bodily

sensations, are parts of the body image. The pain in the foot is literally in the physical space in the brain." (Searle 1992, p63)

However, Searle does not wish to dismiss conscious phenomenology. Indeed, later in the same book, he concludes that

"...consciousness consists in the appearances themselves. Where appearance is concerned we cannot make the appearance-reality distinction because the appearance is the reality." (Searle 1992, p121).

This illustrates the acute problem that apparent spatial location poses for biological naturalism: If biological naturalism is true, experiences are states of the brain, which are necessarily in the brain. However, if "the appearance is the reality", and the pain appears to be in the foot, then it really is in the foot. Either biological naturalism is true, or the appearance is the reality. One can't have both.

Has science discovered that (despite appearances) pains are really in the brain as Searle suggests? It is true of course that science has discovered *representations* of the body in the brain, for example, a tactile mapping of the body surface distributed over the somatosensory cortex (SSC). However, no scientist has observed actual body sensations to be in the brain, and no scientist ever will, for the simple reason that, viewed from an external observer's perspective, the body *as experienced by the subject* cannot be observed (one cannot directly observe another person's experience). Science has nevertheless investigated the *relationship* of the body image (in SSC) to tactile experiences. Penfield & Rassmussen (1950), for example, exposed areas of cortex preparatory to surgical removal of cortical lesions responsible for focal epilepsy. To avoid surgical damage to areas essential to normal functioning, they explored the functions of these areas by lightly stimulating them with a microelectrode and noting the subject's consequent experiences. As expected, stimulation of the somatosensory cortex produced reports of tactile experiences. However, these feelings of numbness, tingling and so on were subjectively located *in different regions of the body, not in the brain.* In sum, science has discovered that neural excitation of somatosensory cortex *causes* tactile sensations, which are subjectively located in different regions of the body. This effect is precisely the "perceptual projection" that the reflexive model describes.

In recent years the spatially extended nature of visual experience has once more become a topical issue. For example, Pribram (1971, 2004), one of the first scientists to address this problem, has continued to develop his earlier theories of holographic representation in the brain; Revonsuo (1995) developed the suggestion that the phenomenal world is a form of virtual reality (see also Velmans, 1993); and Lehar (2003) in a recent *BBS* target article has attempted to develop a mathematical model of how objects *appear* as they move in phenomenal space (as opposed to how they really *are* as they move in phenomenal space). As these, and other scientists (such as Gray, 2004) have pointed out, the 3D nature of the phenomenal world is likely to have important consequences for neuroscience, for the obvious reason that the brain has to be organised in a way that supports such spatially extended experiences.

However, these theorists remain divided on the issue of whether some experiences are really outside the brain. Pribram (2004) takes the view that they are, and outlines a broad theory of perception that he explicitly links to the reflexive model developed in Velmans (2000). Revonsuo, Lehar and Gray adopt a form of biological naturalism, arguing for example that the entire 3D phenomenal world, stretching to the horizon and the dome of the sky, is a form of virtual reality that is literally inside the brain.

Paradigm crunch

Lehar (2003), however, points out that if the phenomenal world is inside the brain, the real skull must be outside the phenomenal world (the former and the latter are logically equivalent). Let me be clear: if one accepts that

a) The phenomenal world appears to have spatial extension to the perceived horizon and dome of the sky.
b) The phenomenal world is really inside the brain.

It follows that

c) The real skull (as opposed to the phenomenal skull) is beyond the perceived horizon and dome of the sky.

Although Lehar accepts this conclusion, he admits that this consequence of biological naturalism is "incredible". In my view, this casts an entirely different light on the so-called 'scientific' status of biological naturalism and the so-called 'unscientific' claims of the reflexive model. Put your hands on your head. Is that the real skull that you feel, located more or less where it seems to be? If that makes sense, the reflexive model makes sense. Or is that just a phenomenal skull inside your brain, with your real skull beyond the dome of the sky? If the latter seems absurd, biological naturalism is absurd. Choose for yourself.[11]

References

Blackmore, S.J. (2002) There is no stream of consciousness. *Journal of Consciousness Studies,* 9 (5-6)**,** 17-28.

Dennett, D (2002) How could I be wrong? How wrong could I be? *Journal of Consciousness Studies*, 9 (5-6), 13-16.

Gray, J. (2004) *Consciousness: Creeping Up On The Hard Problem.* Oxford: Oxford University Press.

Jack, A.I. and Prinz, J.J. (2004) Searching for a scientific experience. Journal of Consciousness Studies, 11(1), 51-55.

Lehar, S. (2003) Gestalt isomorphism and the primacy of subjective conscious experience: A gestalt bubble model. *Behavioral & Brain Sciences* 26(4), 375-444.

Miller, G. and Johnson-Laird, P. (1976) *Language and Perception.* Cambridge: Cambridge University Press.

Noe, A. and Thomson, E. (2004a) Are there neural correlates of consciousness? Journal of Consciousness Studies 11(1), 3-28.

Noe, A. and Thomson, E. (2004b) Sorting out the neural basis of consciousness. Journal of Consciousness Studies, 11(1), 87-98.

O'Regan, J. K., Myin, E. and Noe, A. Towards an analytic phenomenology: the concepts of "bodiliness" and "grabbiness." In A. Carsetti (Ed.) Proceedings of the International Colloquium : "Seeing and Thinking. Reflections on Kanizsa's Studies in Visual Cognition". Univ. Tor Vergata, Rome, June 8-9, 2001. Kluwer (in press).

Penfield, W. and Rassmussen, T.B. (1950) *The Cerebral Cortex of Man*, Princeton: Princeton University Press.

Pribram, K.H. (1971) *Languages of the brain: experimental paradoxes and principles in neuropsychology*. New York: Brandon House.

Pribram, K. (2004) Consciousness reassessed. *Mind and Matter,* 2(1), 7-35.

Revonsuo, A. (1995) Consciousness, dreams, and virtual realities. *Philosophical Psychology*, 8(1): 35-58.

Searle, J. (1992) *The Rediscovery of the Mind*, Cambridge, Mass: MIT Press.

Simons, D.J. and Chabris, C. (1999) Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception*, 28(9), 1059-1074.

Simons D.J and Levin D.T., (1998) Failure to detect changes to people in a real-world interaction" *Psychonomic Bulletin and Review*, 5, 644 – 649.

Velmans, M. (1990) Consciousness, brain, and the physical world. *Philosophical Psychology*, 3, 77-99.

Velmans, M.(1993) A Reflexive Science of consciousness. In *Experimental and Theoretical Studies of Consciousness. CIBA Foundation Symposium 174.* Wiley, Chichester, pp 81-99.

Velmans, M. (1998) Goodbye to reductionism. In S.Hameroff, A.Kaszniak & A.Scott (eds) *Towards a Science of Consciousness II: The Second Tucson Discussions and Debates.* MIT Press, pp 45-52.

Velmans, M. (2000) *Understanding Consciousness*, London: Routledge/Psychology Press.

Velmans, M. (2001). Heterophenomenology versus critical phenomenology: a dialogue with Dan Dennett. Online at http://cogprints.ecs.soton.ac.uk/archive/00001795/index.html

Velmans, M. (2003a) *How Could Conscious Experiences Affect Brains?* Exeter: Imprint Academic.

Velmans, M (2003b) Is the brain in the world, or the world in the brain? *Behavioral and Brain Sciences*, 26(4): 427-429.

[1] For a more detailed critique of Dennett's views on the existence of qualia see Velmans (2001).

[2] The Ames room looks square, but one side of the facing wall is actually much further away than the other with the consequence that a person positioned at the near edge appears much larger than a person at the more distant edge.

[3] See Velmans (1991) for an extensive review.

[4] I am not suggesting that functioning can be *completely* dissociated from accompanying experience in normal human beings. It is unlikely for example, that experiences can be carved off from a normally functioning brain with a surgeon's knife, without in some way disrupting that functioning. But that does not settle the question of how functioning *relates* to the accompanying experience—or even, whether the functioning (enactive or not) somehow *explains* the experience.

[5] See detailed discussion of this and related points in Velmans (2000) chapters 4 and 5.

[6] See for example the classical studies of temporal lobe stimulation carried out by Penfield & Rassmussen (1950).

[7] Although I accept that preconscious perceptual processing may also involve dynamic sensory-motor interactions of the kind suggested by the enactive approach above.

[8] Note, however, that the reflexive model is not externalist (for any doctrinal reason) about all experiences. Whether an experience is located in external phenomenal space, on the body surface, or nowhere, is an empirical matter that is entirely dependent on its phenomenology. For example, the phonemic imagery that accompanies the thought that 2+2=4 does not have a clear location, or might seem, at best, to be roughly located, "inside the head" (see Velmans, 2000, ch6).

[9] There are of course alternative representations of space suggested by physics (four dimensional space time, 11 dimensional space of string theory, etc) and non-Cartesian geometries (e.g. Riemann geometry). A comparison of phenomenal to measured (Cartesian) space is all that we need however to decide whether a pain in my foot or this perceived print on this page is, or is not, really in my brain.

[10] Although it is not germane to the issue under discussion, it should be noted that the reflexive model adopts a form of critical realism, which accepts that the phenomenal world represents an autonomously existing world itself with varying degrees of utility and accuracy—and that phenomenal space, measured space, and the various conceptualisations of space developed in physics are alternative representations of space itself whose utility and accuracy can only be assessed in the light of the purposes for which it is to be used (cf Velmans, 2000, ch7).

[11] For further discussion, see Velmans (2003b).