

Survey Article

Natural theories of consciousness

Peter Carruthers

Many people have thought that consciousness – particularly phenomenal consciousness, or the sort of consciousness which is involved when one undergoes states with a distinctive subjective phenomenology, or ‘feel’ – is inherently, and perhaps irredeemably, mysterious (Nagel 1974, 1986; McGinn 1991). And many would at least agree with Chalmers (1996) in characterising consciousness as the ‘hard problem’, which forms one of the few remaining ‘final frontiers’ for science to conquer. But equally, there have been a plethora of attempts by philosophers at explaining consciousness in natural terms,¹ many of them quite recent (Armstrong 1968, 1984; Carruthers 1996; Dennett 1978, 1991; Dretske 1995; Flanagan 1992; Gennaro 1996; Kirk 1994; Lycan 1987, 1996; Nelkin 1996; Rosenthal 1986, 1993; Tye 1995). This paper surveys the prospects for success of such attempts, focusing particularly on the recent books by Dretske (1995), Tye (1995), Gennaro (1996) and Lycan (1996).² But it is by no means impartial; and the reader should note that I have my own axe to grind in this exercise. My overt agenda is to convince you of the merits of *dispositionalist higher-order thought theories* in particular, of the sort defended in my 1996, chs. 5 to 7.³

1. Some distinctions, and a road-map

One of the real advances made in recent years has been in distinguishing between different questions concerning consciousness (see particularly: Dretske 1993; Block 1995; Lycan). Not everyone agrees on quite *which* distinctions need to be drawn, however; and I shall be arguing later that one crucial distinction (between worldly-subjectivity and mental-state-subjectivity) has been overlooked. But all are agreed that we should distinguish *creature* consciousness from *mental-state* consciousness. It is one thing to say of *an individual person or organism* that it is conscious (either in general or of something in particular); and it is quite another thing to say of *one of the mental states* of a creature that it is conscious.

It is also agreed that within creature-consciousness itself we should distinguish between *intransitive* and *transitive* variants. To say of an organism that it is conscious *simpliciter* (intransitive) is to say just that it is awake, as opposed to

asleep or comatose. There do not appear to be any deep philosophical difficulties lurking here (or at least, they are not difficulties specific to the topic of consciousness, as opposed to mentality in general). But to say of an organism that it is conscious of *such-and-such* (transitive) is normally to say at least that it is *perceiving* such-and-such. So we say of the mouse that it is conscious of the cat outside its hole, in explaining why it does not come out; meaning that it *perceives* the cat's presence. To provide an account of transitive creature-consciousness would thus be to attempt a theory of perception. No doubt there *are* many problems here; but I shall proceed as if I had the solution to them.

There is a choice to be made concerning transitive creature-consciousness, failure to notice which may be a potential source of confusion. For we have to decide whether the perceptual state in virtue of which an organism may be said to be transitively-conscious of something must itself be a conscious one (state-conscious – see below). If we say 'Yes' then we shall need to know more about the mouse than merely that it perceives the cat if we are to be assured that it is conscious of the cat – we shall need to establish that its percept of the cat is itself conscious. If we say 'No', on the other hand, then the mouse's perception of the cat will be sufficient for the mouse to count as conscious of the cat; but we may have to say that although it is conscious of the cat, the mental state in virtue of which it is so conscious is not itself a conscious one! I think it best to by-pass all danger of confusion here by avoiding the language of transitive-creature-consciousness altogether. Nothing of importance would be lost to us by doing this. We can say simply that organism O *observes* or *perceives* X; and we can then assert explicitly, if we wish, that its percept is or is not conscious.

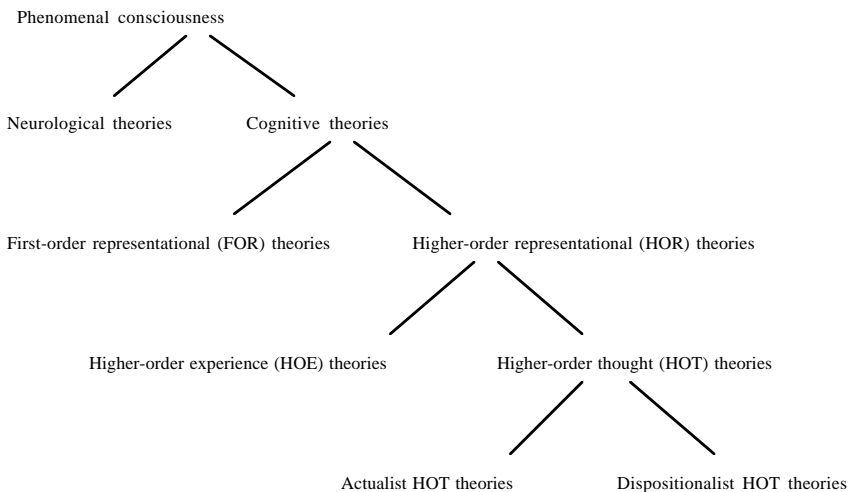
Turning now to the notion of *mental-state consciousness*, the major distinction is between *phenomenal consciousness*, on the one hand – which is a property of states which it is *like something* to be in, which have a distinctive 'feel' – and various functionally-definable notions, such as Block's (1995) *access consciousness*, on the other. Most theorists believe that there are mental states – such as occurrent thoughts or judgements – which are conscious (in whatever is the correct functionally-definable sense), but which are not phenomenally conscious.⁴ But there is considerable dispute as to whether mental states can be phenomenally-conscious without also being conscious in the functionally-definable sense – and even more dispute about whether phenomenal consciousness can be *explained* in functional and/or representational terms.

It seems plain that there is nothing deeply problematic about functionally-definable notions of mental-state consciousness, from a naturalistic perspective. For mental functions and mental representations are the staple fare of naturalistic accounts of the mind. But this leaves plenty of room for dispute about the form that the correct functional account should take. Some claim that for a state to be conscious in the relevant sense is for it to be poised to have an impact on the organism's decision-making processes (Kirk 1994; Dretske; Tye), perhaps also with the additional requirement that those processes should be distinctively rational ones (Block 1995). Others think that the relevant requirement is that the state should be suitably related to higher-order representations – beliefs and/or

experiences – of that very state (Armstrong 1984; Dennett 1991; Rosenthal 1993; Carruthers 1996; Gennaro; Lycan).

What is often thought to be naturalistically problematic, in contrast, is phenomenal consciousness (Nagel 1984; McGinn 1991; Block 1995; Chalmers 1996). And what is really and deeply controversial is whether phenomenal consciousness can be *explained* in terms of some or other functionally-definable notion. *Cognitive* theories maintain that it can – the contrast here being with those who think that phenomenal consciousness should be explained in neurological terms, say in terms of oscillation-patterns amongst neural spiking-frequencies (e.g. Crick and Koch 1990).

Naturalistic theories of phenomenal consciousness may then usefully be sorted along a series of choice-points, as represented in the diagram below. First, the theorist has to decide whether to try and account for consciousness in *physical* and/or *neurological*, or rather in *cognitive* (representational and/or functional) terms.⁵ Here I shall just assume that the correct choice is the latter. We begin our discussion at the second choice-point: between theories which account for phenomenal consciousness in terms purely of first-order representations (FORs) of the environment (or of the subject's own body); and theories which involve higher-order representations (HORs) of the subject's own mental states. I shall argue that we ought to go for the right-hand branch. Then the third choice-point is between *inner sense* models of consciousness, which conceive of the higher-order representations which render a given mental state conscious as being somewhat like *experience* (that is, higher-order experience, or HOE, models); and higher-order thought (HOT) theories, which explain consciousness in terms of thoughts or judgements about our mental states. Again I defend the right-hand branch. Then finally the choice is between accounts which require the *actual* presence of a HOT for a state to qualify as conscious; and accounts which characterise



conscious states as those which are suitably *available* to HOTs. Once again I argue that the right-hand branch is to be preferred.

2. FOR theories

In two recent, wonderfully written, lucid, and highly ambitious books, Dretske and Tye have independently developed very similar first-order representationalist (FOR) theories of phenomenal consciousness. In both cases the goal is to characterise all of the phenomenal – ‘felt’ – properties of experience in terms of the representational *contents* of experience. So the difference between an experience of green and an experience of red will be explained as a difference in the properties represented – reflective properties of surfaces, say – in each case. And the difference between a pain and a tickle is similarly explained in representational terms – the difference is said to reside in the different properties (different kinds of disturbance) represented as located in particular regions of the subject’s own body. In each case, a phenomenally-conscious experience is said to be one which is *poised* to have an impact on the subject’s beliefs and practical-reasoning processes in such a way as to guide behaviour.

Perhaps the main consideration supporting such FOR-theories is the so-called ‘transparency’ of consciousness, previously noted by a number of writers (e.g. Harman 1990; McCulloch 1988, 1993). Look at a green tree or a red tomato. Now try to concentrate as hard as you can, *not* on the colours of the *objects*, but on the quality of your *experience* of those colours. What happens? Can you do it? Plausibly, all that you find yourself doing is paying closer and closer attention to the colours in the outside world, after all. A perception of red is a state which represents a surface as having a certain distinctive quality – *redness*, of some or other particular shade – and paying close attention to your perceptual state comes down to paying close attention to the quality of the world *represented* (while being aware of it *as* represented – this will become important later). Of course, in cases of perceptual illusion or hallucination there may actually be no real quality of the world represented, but only a *representing*. But still, plausibly, there is nothing to your experience over and above the way it represents the world as being.

But what about bodily sensations, like itches, tickles, and pains? Are these, too, purely representational states? If so, *what* do they represent? It might seem that all there really is to a pain, is a particular sort of *non-representational* quality, which is experienced as unwelcome. But in fact, as Tye shows, the case for non-representational qualities (*qualia*) is no stronger in connection with pain than it is with colour. In both cases our experience represents to us a particular perceptible property – in the one case, of an external surface, in the other case, of a region of our own body. In the case of colour-perception, my perceptual state delivers the content, [*that* surface has *that* quality]. In the case of pain, my state grounds an exactly parallel sort of content, namely [*that* region of my body has *that* quality]. In each case the *that quality* expresses a recognitional concept, where what is

recognised is not a quale, but rather a property which our perceptual state represents as being instantiated in the place in question.

Dretske and Tye differ from one another mainly in the accounts that they offer of the representation-relation. For Dretske, the content of a representational state is fixed teleologically, in terms of the objects/properties which that state is *supposed* to represent, given the organism's evolutionary and learning histories. For Tye, in contrast, the content of a state is defined in terms of causal co-variance in normal circumstances – where the notion of *normal* circumstances may or may not be defined teleologically, depending upon cases. But both are agreed that content is to be individuated *externally*, in a way which embraces objects and properties in the organism's environment. I begin my discussion by suggesting that they have missed a trick in going for an externalist notion of content, and that their position would be strengthened if they were to endorse a *narrow*-content account instead.

Consider the now-famous example of *Swampman* (Davidson 1987), who is accidentally created by a bolt of lightning striking a tree-stump in a swamp, in such a way as to be molecule-for-molecule identical to Davidson himself. Dretske is forced to deny that Swampman and Davidson are subject to the same colour-experiences (and indeed, he must deny that Swampman has any colour-experiences at all), since his states lack *functions*, either evolved or learned. As Dretske admits, this consequence is highly counter-intuitive; and it is one which he has to chew on pretty hard to force himself to swallow. Tye, on the other hand, believes that he is better off in relation to this example, since he says that Swampman's circumstances can count as 'normal' by default. But then there will be other cases where Tye will be forced to say that two individuals undergo the same experiences (because their states are such as to co-vary with the same properties in circumstances which are normal for them), where intuition would strongly suggest that their experiences are different.

So, imagine that the lightning-bolt happens to create Swampman with a pair of colour-inverting spectacles permanently attached to his nose. Then Tye may have to say that when Swampman looks at green grass, he undergoes the same experiences as Davidson does (who views the grass *without* such glasses). For in the circumstances which are normal for Swampman, he is in a state which will co-vary with greenness. So he experiences *green*, just as Davidson does. This, too, is highly counter-intuitive. We would want to say, surely, that Swampman experiences *red*.

Tye (in correspondence) replies that by *normal* he just means *ceteris paribus*, and that if Swampman has colour-inverting spectacles in front of his eyes, then *ceteris* is not really *paribus*. But now it looks easy to develop the example in such a way as to re-instate the point. Suppose that the bolt of lightning creates colour-inverting lenses as part of the very structure of the cornea within Swampman's eyes. Then other things are, surely, equal for *Swampman* when he looks at green grass, in which case Tye's sort of FOR-theory will have to say that Swampman has an experience *as of* green; but again, intuition strongly suggests that his experiences will be the inverse of Davidson's.

Some may see sufficient reason, here, to reject a FOR-account of phenomenal consciousness straight off. I disagree. Rather, such examples just motivate adoption of a narrow-content account of representation in general, where contents are individuated in abstraction from the particular objects and properties in the thinker's environment. We might say, for example, that the representational content of the experience is the same whenever it is such that it *would* engage the very same recognitional capacity (note that this need not imply that the content of the experience is itself conceptualised). But explaining and defending such an account, and showing how it is fully consistent with – indeed, required by – naturalism about the mental, would take me too far afield for present purposes.⁶ In the present context I must ask readers just to bracket any worries they may have about externalist theories of perceptual content, accepting on trust that one could embrace a FOR-naturalisation of phenomenal consciousness while rejecting externalism.

This is not to say, of course, that I think FOR-approaches to consciousness are unproblematic. On the contrary: over the next two sections I shall develop objections which seem to me to count decisively in favour of some sort of HOR-approach.

3. First problem: phenomenal world versus phenomenal experience

One major difficulty with FOR-accounts in general, is that they cannot distinguish between what the *world* (or the state of the organism's own body) is like for an organism, and what the organism's *experience of the world* (or of its own body) is like for the organism. This distinction is very frequently overlooked in discussions of consciousness. And Tye, for example, will move (sometimes in the space of a single sentence) from saying that his account explains what *colour* is like for an organism with colour-vision, to saying that it explains what *experiences of colour* are like for that organism. But the first is a property of the world (or of a world-perceiver pair, perhaps), whereas the latter is a property of the organism's experience of the world (or of an experience-experencer pair). These are plainly distinct.

It is commonplace to note that each type of organism will occupy a distinctive point of view on the world, characterised by the kinds of perceptual information which are available to it, and by the kinds of perceptual discriminations which it is capable of making (Nagel 1974). This is part of what it means to say that bats (with echolocation) and cats (without colour vision) occupy a different point of view on the world from ourselves. Put differently but equivalently: the world (including subjects' own bodies) is *subjectively presented* to different species of organism somewhat differently. And to try to characterise this is to try and understand what the world for such subjects *is like*. But it is one thing to say that *the world* takes on a subjective aspect by being presented to subjects with differing conceptual and discriminatory powers, and it is quite another thing to say that the subject's *experience of the world* also has such a subjective

aspect, or that there is something which the *experience* is *like*. Indeed, by parity of reasoning, this would seem to require subjects to possess information about, and to make discriminations amongst, their own states of experience. And it is just this which provides the rationale for HOR-accounts as against FOR-accounts, in fact.

According to HOR-theories, first-order perceptual states (if non-conscious – see Section 4 below) may be adequately accounted for in FOR terms. The result will be an account of the point of view – the subjective perspective – which the organism takes towards its world (and the states of its own body), giving us an account of what the world, for that organism, *is like*. But the HOR-theorist maintains that something else is required in accounting for what an *experience* is like for a subject, or in explaining what it is for an organism's *mental states* to take on a subjective aspect. For this, we maintain, higher-order representations – states which meta-represent the subject's own mental states – are required. And it is hard to see how it could be otherwise, given the distinction between what the world is like for an organism, and what its experience of the world is like.

We therefore need to distinguish between two different sorts of subjectivity – between worldly-subjectivity and mental-state-subjectivity. In fact we need to distinguish between phenomenal properties of the world (or of the organism's own body), on the one hand, and phenomenal properties of one's experience of the world (or of one's experience of one's body) on the other. FOR-theory may be adequate to account for the former; but not to explain the latter, where some sort of HOR-theory is surely needed. Which of these two deserves the title 'phenomenal consciousness'? There is nothing (or nothing much) in a name; and I am happy whichever reply is given. But it is the subjectivity of experience which seems to be especially problematic – if there is a 'hard problem' of consciousness (Chalmers 1996), it surely lies here. At any rate, nothing can count as a complete theory of phenomenal consciousness which cannot explain it – as FOR-theory plainly cannot.

But now, how is HOR-theory to handle the point that conscious experiences have the quality of *transparency*? As we noted earlier, if you try to focus your attention on your experience of a bright shade of colour, say, what you find yourself doing is focusing harder and harder on the colour itself – your focus seems to go right *through* the experience to its objects. This might seem to lend powerful support to FOR-accounts of phenomenal consciousness. For how can any form of HOR-theory be correct, given the transparency of experience, and given that all the *phenomena* in phenomenal consciousness seem to lie in *what is represented*, rather than in anything to do with the mode of representing it?

Now in one way this line of thought is correct – for in one sense there is nothing in the content of phenomenally-conscious experience beyond what a FOR-theorist would recognise. What gets added by the presence of a HOR-system is a dimension of *seeming* or *appearance* of that very same first-order content. But in another sense this *is* a difference of content, since the content *seeming red* is distinct from the content *red*. So when I focus on my experience of a colour I can,

in a sense, do other than focus on the colour itself – I can focus on the way that colour *seems* to me, or on the way it *appears*; and this is to focus on the subjectivity of my experiential state. It is then open to us to claim that it is the possibility of just such a manner of focusing which confers on our experiences the dimension of subjectivity, and so which renders them for the first time fully phenomenally-conscious, as we shall see in Section 5.

4. Second problem: conscious *versus* non-conscious experience

Another – closely related – difficulty with FOR-approaches is to provide an account of the distinction between conscious and non-conscious experience. (As examples of the latter, consider absent-minded driving; sleepwalking; experience during mild epileptic seizure; and blindsight.) For in some of these cases, at least, we appear to have first-order representations of the environment which are not only poised for the guidance of behaviour, but which are actually controlling it.⁷ So how can FOR-theorists explain why our perceptions, in such cases, are *not* phenomenally-conscious? There would seem to be just two ways for them to respond – either they can accept that absent-minded driving experiences are not phenomenally-conscious, and characterise what additionally is required to render an experience phenomenally-conscious in (first-order) functional terms; or they can insist that absent-minded driving experiences *are* phenomenally-conscious, but in a way which makes them inaccessible to their subjects.

Kirk (1994) apparently exemplifies the first approach, claiming that for a perceptual state with a given content to be phenomenally conscious, and to acquire a 'feel', it must be present to the *right sorts* of decision-making processes – namely those which constitute the organism's highest-level executive. But this is extremely puzzling. It is utterly mysterious how an experience with one and the same content could be sometimes phenomenally-conscious and sometimes not, depending just upon the overall role in the organism's cognition of the decision-making processes to which it is present.⁸

Tye takes the second approach. In cases such as that of absent-minded driving, he claims that there is experience, which is phenomenally-conscious, but which is 'concealed from the subject'.⁹ This then gives rise to the highly counter-intuitive claim that there are phenomenally-conscious experiences to which the subject is blind – experiences which it is *like* something to have, but of which the subject is unaware. And in explaining the aware/unaware distinction, Tye then goes for an actualist form of HOT-theory. He argues that we are *aware* of an experience and its phenomenal properties only when we are actually applying phenomenal concepts to it. The dilemma then facing him, is *either* that he cannot account for the immense richness of experience of which we are (can be) aware; *or* that he has to postulate an immensely rich set of HOTs involving phenomenal concepts accompanying each set of experiences of which we are aware – the same dilemma faced by any actualist HOT-theorist, in fact (see Section 7 below).

Not only is Tye's position counter-intuitive, but it may actually be incoherent. For the idea of the *what-it-is-likeness* of *experience* is intended to characterise those aspects of experience which are subjective. But there surely could not be properties of experience which were subjective without being *available* to the subject, and of which the subject was unaware. An experience of which the subject is unaware cannot be one which it is *like something* for the *subject* to have. On the contrary, an experience which it is *like something* to have, must be one which is available to the subject of that experience – and that means being a target (actual or potential) of a suitable HOR.

It may be objected that 'subjective' just implies 'grounded in properties of the subject', and that we are quite happy with the idea that someone's prejudices, for example, might reflect evaluations which are subjective in this sense without being available to the subject. But in the case of perception it is already true that the subjectivity of the world, for a subject, is grounded in properties of the perceiver. In what could the further subjectivity of *the experience* consist, except its availability to the subject? Since the way the world appears – subjectively – to a subject depends upon the way properties of the world are made available to the subject, grounded in properties of the subject's perceptual system, it is hard to see in what else the subjectivity of the subject's experience of the world could consist but *its* availability to the subject in turn, through some type of HOR.

5. The explanatory power of HOR theories

I now propose to argue that phenomenal consciousness will emerge, of natural necessity (perhaps also metaphysical: I shall not pursue this here), in any system where perceptual information is made available to HORs in analogue form, and where the system is capable of recognising its own perceptual states, as well as the states of the world perceived. For by postulating that this is so, we can explain why phenomenal feelings should be so widely thought to possess the properties of *qualia* – that is, of being non-relationally defined, private, ineffable, and knowable with complete certainty by the subject.¹⁰ I claim, in fact, that any subjects who instantiate such a cognitive system (that is, who instantiate a HOR-model of state-consciousness) will normally come to form just such beliefs about the intrinsic characteristics of their perceptual states – and they will form such beliefs, not because they have been explicitly programmed to do so, but naturally, as a by-product of the way in which their cognition is structured. This then demonstrates, I believe, that a regular capacity for HORs about one's own mental states must be a sufficient condition for the enjoyment of experiences which possess a subjective, phenomenal, feel to them.¹¹

Let us consider, in particular, the thesis of non-relational definition for terms referring to the subjective aspects of an experience. This is a thesis which many people find tempting, at least. When we reflect on what is essential for an experience to count as an experience *as of red*, for example, we are inclined to deny

that it has anything directly to do with being caused by the presence of something red. We want to insist that it is conceptually possible that an experience of that very type should normally have been caused by the presence of something green, say. All that is truly essential to the occurrence of an experience *as of red*, on this view, is the way such an experience feels to us when we have it – it is the distinctive feel of an experience which defines it, not its distinctive relational properties or causal role (see Kripke 1972).

Now any system instantiating a HOR-model of consciousness will have the capacity to distinguish or classify perceptual states according to their contents, not by inference (that is, by self-interpretation) or description, but immediately. The system will be capable of recognising the fact that it has an experience *as of red*, say, in just the same direct, non-inferential, way that it can recognise red. (This is just what it means to say that perceptual states are available to HORs, in the intended sense.) The system will, therefore, readily have available to it purely recognitional concepts of experience. In which case, absent and inverted subjective feelings will immediately be a conceptual possibility for someone applying these recognitional concepts. If I instantiate such a system, I shall straight away be able to think, '*This type of experience might have had some quite other cause*', for example.

I have conceded that there may be concepts of experience which are purely recognitional, and so which are not definable in relational terms. Does this then count against the acceptability of the functionalist conceptual scheme which forms the background to cognitive accounts of consciousness? If it is conceptually possible that an experience *as of red* should regularly be caused by perception of green grass or blue sky, then does this mean that the crucial facts of consciousness must escape the functionalist net, as many have alleged? I think not. For HOR-accounts are not in the business of conceptual analysis, but of substantive theory development. So it is no objection to those accounts, that there are some concepts of the mental which cannot be analysed (that is, defined) in terms of functional or representational role, but are purely recognitional – provided that the nature of those concepts, and the states which they recognise, can be adequately characterised within the theory.

According to HOR-theory, the properties which are in fact picked out (note: not *as such*) by any purely-recognitional concepts of experience are not, themselves, similarly simple and non-relational.¹² When I recognise in myself an experience *as of red*, what I recognise is, in fact, a perceptual state which represents worldly redness, and which underpins, in turn, my capacity to recognise, and to act differentially upon, red objects. And the purely-recognitional concept, itself, is one which represents the presence of just such a perceptual state, and tokenings of that concept then cause further characteristic changes within my cognition. There is nothing, here, which need raise any sort of threat to a naturalistic theory of the mind.

With the distinction firmly drawn between our recognitional *concepts* of phenomenal feelings, on the one hand, and the *properties* which those concepts pick out, on the other, we can then claim that it is naturally (perhaps

metaphysically) necessary that the subjective aspect of an experience of red should be caused, normally, by perception of red. For a HOR-account tells us that the subjective aspect of an experience of red just *is* an analogue representation of red, presented to a cognitive apparatus having the power to distinguish amongst states in terms of their differing representational contents, as well as to classify and distinguish between the items represented. In which case there can be no world (where the laws of nature remain as they are, at least) in which the one exists but not the other. For there will, in fact, be no 'one' and 'other' here, but only one state differently thought of – now recognitionally, now in terms of functional role.

But is it not possible – naturally and metaphysically, as well as conceptually – that there should be organisms possessing colour vision, which are sensitive to the same range of wavelengths as ourselves, but which nevertheless have their phenomenal feelings inverted from ours? I claim not, in fact. For the property of being the subjective feel of an experience of red is a functional/representational one, identical with possession of a distinctive causal/representational role (the role, namely, of being a state which represents worldly redness, and which is present to a HOR-faculty with the power to recognise its own perceptual states as such). In which case feeling-inversion of the type imagined will be impossible.

Since any organism instantiating a HOR-model of state-consciousness will naturally be inclined to make just those claims about its experiences which human *qualia-freaks* make about theirs, we have good reason to think that HOR-theory provides us with a *sufficient* condition of phenomenal consciousness. But is there any reason to think that it is also *necessary* – that is, for believing that HOR-theory gives us the truth about what phenomenal consciousness *is*? One reason for doubt is that a FOR-theorist, too, can avail himself of the above explanation (as Tye does, for example). For FOR-theorists need not deny that we humans are in fact capable of HORs. They can then claim that FOR-theory gives the truth about phenomenal consciousness, while appealing to HORs to explain, e.g., the conceptual possibility of inverted spectra. To put the point somewhat differently: it may be claimed that what *underpins* the possibility of inverted spectra (i.e. phenomenal consciousness itself) is there, latent, in FOR-systems; but that only a creature with the requisite concepts (HORs) can actually *entertain* that possibility.

This suggestion can be seen to be false, however, in light of the FOR-theorists' failure to distinguish between worldly-subjectivity and mental-state-subjectivity, discussed in Section 3 above. In fact a system which is capable only of FORs will have the raw-materials to underpin only a much more limited kind of possibility. Such a system may contain, let us say, FORs of *red*. Its states will then represent various surfaces as covered with a certain uniform property, for which it may possess a recognitional concept. This provides the raw materials for thoughts such as, '*That property [red] may in fact be such-and-such a property [e.g. pertaining to reflective powers]*'. But there is nothing here which might make it possible to entertain thoughts about spectra inversion. Lacking any way of distinguishing between *red* and *the experience of red*, the system lacks the raw-materials necessary

to underpin such thoughts as, 'Others may experience *red* as I experience *green*' – by which I mean not just that a FOR-system will lack the concepts necessary to frame such a thought (this is obvious), but that there will be nothing *in the contents of the system's experiences and other mental states* which might warrant it.

6. Conscious states for animals (and young children)?

Having argued for the superiority of HOR-theory over FOR-theory, I turn now to the question of how widely distributed conscious mental states will be, on a HOR-account. For both Dretske and Tye claim – without any real argument – that this provides a decisive consideration in favour of their more modest FOR-approach. I shall argue that they are right to claim that HOR-theories must deny phenomenal-consciousness to the mental states of animals (and very young children), but wrong that this provides any reason for accepting a FOR-account.

Gennaro defends a form of HOT-theory. And he acknowledges that if possession of a conscious mental state *M* requires a creature to conceptualise (and entertain a HOT about) *M as M*, then probably very few creatures besides human beings will count as having conscious states. Let us focus on the case where *M* is a percept of green, in particular. If a conscious perception of a surface as green required a creature to entertain the HOT [that I am perceiving a green surface], then probably few other creatures, if any, would qualify as subjects of such a state. There is intense debate about whether even chimpanzees have a conception of perceptual states as such (see, e.g., Povinelli 1996); in which case it seems very unlikely that any non-apes will have one. So the upshot might be that phenomenal-consciousness is restricted to apes, if not exclusively to human beings.

This is a consequence which Gennaro is keen to resist. He tries to argue that much less conceptual sophistication than the above is required. In order for *M* to count as conscious one does not have to be capable of entertaining a thought about *M qua M*. It might be enough, he thinks, if one were capable of thinking of *M as distinct from* some other state *N*. Perhaps the relevant HOT takes the form, [*this is distinct from that*]. This certainly *appears* to be a good deal less sophisticated. But appearances can be deceptive – and in this case I believe that they are.

What would be required in order for a creature to think, of an experience of green, that it is distinct from a concurrent experience of red? More than is required for the creature to think *of green* that it is distinct from red, plainly – this would not be a HOT at all, but rather a first-order thought about the distinctness of two perceptually-presented colours. So if the subject thinks, [*this is distinct from that*], and thinks something higher-order thereby, then *something* must make it the case that the relevant *this* and *that* are colour experiences as opposed to just colours. What could this be?

There would seem to be just two possibilities. Either, on the one hand, the *this* and *that* are picked out as experiences by virtue of the subject deploying – at least covertly – a concept of *experience*, or some near equivalent (such as a concept of *seeming*, or *sensation*, or some narrower versions thereof, such as *seeming colour* or

seeming red). This would be like the first-order case where I entertain the thought, [*that is dangerous*], in fact thinking about a particular perceptually-presented cat, by virtue of a covert employment of the concept *cat*, or *animal*, or *living thing*. But this first option just returns us to the view that HOTs (and so phenomenal consciousness) require possession of concepts which it would be implausible to ascribe to most species of animal.

On the other hand, the subject's indexical thought about their experience might be grounded in a non-conceptual *discrimination* of that experience as such. We might model this on the sort of first-order case where someone – perhaps a young child – thinks, [*that is interesting*], of what is in fact a coloured marble (but without possessing the concepts *marble*, *sphere*, or even *physical object*), by virtue of their experience presenting them with a non-conceptual array of surfaces and shapes in space, in which the marble is picked out as one region-of-filled-space amongst others. Taking this second option would move us, in effect, to a *higher-order experience* (HOE) account of consciousness. Just such a view has been defended recently by Lycan, following Armstrong (1968, 1984).¹³

How plausible is it that animals might be capable of HOEs? Lycan faces this question, arguing that HOEs might be widespread in the animal kingdom, perhaps serving to integrate the animal's first-order experiences for purposes of more efficient behaviour-control. But a number of things go wrong here. One is that Lycan seriously underestimates the computational complexity required of the internal monitors necessary to generate the requisite HOEs. In order to perceive an experience, the organism would have to have the mechanisms to generate a set of internal representations with a content (albeit non-conceptual) representing the content of that experience. For remember that both HOT and HOE accounts are in the business of explaining how it is that one aspect of someone's experiences (e.g. of movement) can be conscious while another aspect (e.g. of colour) can be non-conscious. So in each case a HOE would have to be constructed which represents just those aspects, in all of their richness and detail. But when one reflects on the immense computational resources which are devoted to perceptual processing in most organisms, it becomes very implausible that such complexity should be replicated, to any significant degree, in generating HOEs.

Lycan also goes wrong, surely, in his characterisation of what HOEs are *for* (and so, implicitly, in his account of what would have led them to evolve). For there is no reason to think that *perceptual integration* – that is, first-order integration of different representations of one's environment or body – either requires, or could be effected by, second-order processing. So far as I am aware, no cognitive scientist working on the so-called 'binding problem' (the problem of explaining how representations of objects and representations of colour, say, get bound together into a representation of an object-possessing-a-colour) believes that second-order processing plays any part in the process.

Notice, too, that it is certainly not enough, for a representation to count as a HOE, that it should occur down-stream of, and be differentially caused by, a first-order experience. So the mere existence of different stages and levels of

perceptual processing is not enough to establish the presence of HOEs. Rather, those later representations would have to have an appropriate cognitive role – figuring in inferences or grounding judgements in a manner distinctive of second-order representations. What could this cognitive role possibly be? It is very hard to see any other alternative than that the representations in question would need to be able to ground judgements of *appearance*, or of *seeming*, helping the organism to negotiate the distinction between appearance and reality (see my 1996, ch. 5). But that then returns us to the idea that any organism capable of mental-state-consciousness would need to possess *concepts* of experience, and so be capable of HOTs.

I conclude that HOR-theories will entail (when supplemented by plausible empirical claims about the representational powers of non-human animals) that very few animals besides ourselves are subject to phenomenally-conscious mental states.¹⁴ Is this a decisive – or indeed any – consideration in favour of FOR-accounts? My view is that it is not, since we lack any grounds for believing that animals have phenomenally-conscious states. Of course, most of us do have a powerful intuitive belief that there is something which it is *like* for a cat or a rat to experience the smell of cheese. But this intuition is easily explained. For when we ascribe an experience to the cat we quite naturally (almost habitually) try to form a first-person representation of its content, trying to imagine what it might be like ‘from the inside’.¹⁵ But when we do this what we do, of course, is imagine a *conscious* experience – what we do, in effect, is represent one of our *own* experiences, which will bring its distinctive phenomenology with it. All we really have reason to suppose, in fact, is that the cat *perceives* the smell of the cheese. We have no independent grounds for thinking that its percepts will be phenomenally-conscious ones. (Certainly such grounds are not provided by the need to explain the cat’s behaviour. For this purpose the concept of perception, *simpliciter*, will do perfectly well.)

Notice that it is not only animals, but also young children, who will lack phenomenal consciousness according to HOT-accounts. For the evidence is that children under, say, the age of three¹⁶ lack the concepts of *appearance* or *seeming* – or equivalently, they lack the idea of perception as involving *subjective* states of the perceiver – which are necessary for the child to entertain HOTs about its experiences. Dretske uses this point to raise an objection against HOT-theories, which is distinct from the argument from animals discussed above. He asks whether it is not very implausible that three-year-olds and younger children should undergo different *kinds* of experiences – namely, ones which are phenomenally conscious and ones which are not. Granted, the one set of children may be capable of more sophisticated (and higher-order) thoughts than the other; but surely their experiences are likely to be fundamentally the same?

In reply, we may allow that the *contents* of the two sets of experiences are very likely identical; the difference being that the experiences of the younger children will lack the dimension of *subjectivity*. Put differently: *the world* as experienced by the two sets of children will be the same, but the younger children will be blind to the existence and nature of their own experiences. This looks like a pretty

fundamental difference in the mode in which their experiences figure in cognition! – Fundamental enough to justify claiming that the experiences of the one set of children are phenomenally conscious while those of the other are not, indeed.

7. HOE versus HOT, and actualist *versus* dispositionalist

With the superiority of HOR over FOR accounts of phenomenal consciousness now established, the dispute amongst the different forms of HOR-theory is apt to seem like a local family squabble. Accordingly, this section will be brisk.¹⁷

The main problem for HOE-theories, as opposed to HOT-theories, is the problem of *function* – one wonders what all this re-representing is *for*, and how it could have evolved, unless the creature were already capable of entertaining HOTs. In fact this point has already emerged in our discussion of Lycan above – a capacity for higher-order discriminations amongst one's own experiences could not have evolved to aid first-order perceptual integration and discrimination, for example. (Yet as a complex system it surely would have had to evolve, rather than appearing by accident or as an epiphenomenon of some other selected-for function. The idea that we might possess a faculty of 'inner sense' which was not selected for in evolution is surely almost as absurd as the suggestion that *vision* was not selected for – and that is an hypothesis which no one could now seriously maintain.) It is hard to see what function HOEs could serve, in fact, but that of underpinning, and helping the organism to negotiate, the distinction between *appearance* and *reality*. But this is already to presuppose that the creature is capable of HOTs, entertaining thoughts about its own experiences (i.e. about the way things *seem*). And then a creature capable of HOTs wouldn't *need* HOEs – it could just apply its mentalistic concepts directly to, and in the presence of, its first-order experiences (see below).

In contrast, there is no problem whatever in explaining (at least in outline) how a capacity for HOTs might have evolved. Here we can just plug-in the standard story from the primatology and 'theory-of-mind' literatures (see, e.g., Humphrey 1986; Baron-Cohen 1995) – humans might have evolved a capacity for HOTs because of the role such thoughts play in predicting and explaining, and hence in manipulating and directing, the behaviours of others. And once the capacity to think and reason about the beliefs, desires, intentions, and experiences of others was in place, it would have been but a small step to turn that capacity upon oneself, developing recognitional concepts for at least some of the items in question. This would have brought yet further benefits, not only by enabling us to negotiate the appearance/reality distinction, but also by enabling us to gain a measure of control over our own mental lives – once we had the power to recognise and reflect on our own patterns of thought, we also had the power (at least to a limited degree) to change and improve on those patterns; so consciousness breeds cognitive flexibility and improvement.

The main problem for *actualist* as opposed to *dispositionalist* HOT-theories (and note that this is a problem infecting HOE-theories, too, which are also actualist),

is that of *cognitive overload*. There would appear to be an immense amount which we can experience consciously at any one time – think of listening intently to a performance of Beethoven's Seventh Symphony whilst watching the orchestra, for example. But there may be an equally large amount which we can experience *non-consciously*; and the boundaries between the two sets of experiences seem unlikely to be fixed. As I walk down the street, for example, different aspects of my perceptions may be, now conscious, now non-conscious, depending upon my interests, current thoughts, and saliciencies in the environment. Actualist HOR-theories purport to explain this distinction in terms of the presence, or absence, of a HOR targeted on the percept in question. But then it looks as if our HORs must be just as rich and complex as our conscious perceptions, since it is to be the presence of a HOR which explains, of each aspect of those perceptions, its conscious status. And when one reflects on the amount of cognitive space and effort devoted to first-order perception, it becomes hard to believe that a significant proportion of that cognitive load should be replicated again in the form of HORs to underpin consciousness.

The only remotely acceptable response for an actualist HOR-theorist, would be to join Dennett (1991) in denying the richness and complexity of conscious experience.¹⁸ But this is not really very plausible. It may be true that we can only (consciously) *think* one thing at a time (give or take a bit). But there is surely not the same kind of limit on the amount we can consciously *experience* at a time. Even if we allow that a variety of kinds of evidence demonstrates that the periphery of the visual field lacks the kind of determinacy we intuitively believe it to have, for example, there remains the complexity of focal vision, which far outstrips any powers of description we might have.

Dispositionalist forms of HOT-theory can neatly avoid the cognitive overload problem. They merely have to postulate a special-purpose short-term memory store whose function is, *inter alia*, to make its contents available to HOT. The entire contents of the store – which can, in principle, be as rich and complex as you please – can then be conscious in the absence even of a single HOT, provided that the subject remains *capable* of entertaining HOTs about any aspect of its contents. And note that the contents of the store are just first-order percepts, which can then be the objects of HOT – no re-representation is needed.

It is easy to see how a system with the required structure might have evolved. Start with a system capable of first-order perception, ideally with a short-term integrated perceptual memory-store whose function is to present its contents, *poised*, available for use by various theoretical and practical reasoning systems. Then add to the system a 'theory-of-mind' faculty with a capacity for HOTs, which can take inputs from the perceptual memory store, and allow it to acquire recognitional concepts to be applied to the perceptual states and contents of that store. And then you have it! Each of these stages looks like it could be independently explained and motivated in evolutionary terms. And there is minimal meta-representational complexity involved.¹⁹

But is not dispositionalism the wrong *form* for a theory of phenomenal consciousness to take? Surely the phenomenally-conscious status of any given

percept is an *actual* – categorical – property of it, not to be analysed by saying that the percept in question *would* give rise to a targeted HOT in suitable circumstances. In fact there is no real difficulty here. For presumably the percept is *really* – actually – contained in the short-term memory store in question. So the percept is categorically conscious even in the absence of a targeted HOT, by virtue of its presence in the store. It is merely that what constitutes the store as one whose contents are conscious lies in its availability-relation to HOTs.

8. Conclusion

What are the prospects for a naturalistic theory of phenomenal-consciousness? Pretty good, I say. Even FOR-theories have the resources to reply to many of those who think that consciousness is essentially problematic, as Dretske and Tye have shown. And HOR-theories – particularly some form of dispositionalist HOT-theory – can do even better on a number of fronts. It turns out that the ‘hard problem’ is not really so very hard after all!²⁰

Peter Carruthers
 Department of Philosophy
 University of Sheffield
 Sheffield, S10 2TN
 England
 p.carruthers@sheffield.ac.uk

NOTES

¹ I leave to one side in this exercise the many attempts by psychologists at providing an account of consciousness, for simplicity only – my view is actually that there is no sharp line to be drawn between philosophical theories and psychological ones, in this area; both sets of theories are largely intended to be substantive and explanatory, with few *a priori* elements.

² Unless otherwise indicated, references to these four authors are to these works. For comment on Dennett and on Rosenthal, see my 1996, ch. 6. For comment on Nelkin, see my 1997.

³ In fact my main focus in those chapters was on the structure of *human* consciousness, since I was proposing to argue (1996, ch. 8) that natural language is crucially involved in human conscious thinking. In developing an account of consciousness *as such* I would now be inclined to drop the requirement that the higher order thoughts, in virtue of availability to which a given mental state is conscious, must themselves be conscious ones. I am grateful to Colin Allen for helping me to get clear on this.

⁴ In my 1996, ch. 8, I disagreed – arguing that occurrent propositional thoughts can only be conscious (in the human case at least) by being tokened in imaged natural language sentences, which will then possess phenomenal properties.

⁵ *Cognitive* here need not mean *conceptual*. Both FOR-theories (Dretske, Tye) and HOE-

theories (Lycan) maintain that the phenomenal-consciousness-making feature of a mental state consists in a certain sort of *non-conceptual* content.

⁶ See Botterill and Carruthers 1999, ch. 7

⁷ Blindsight may be different, as Tye points out, since in this case there is no behaviour without prompting. Even imagined cases of Super-blindsight (Block 1995) – where subjects become self-cueing, and act spontaneously on information gleaned from their blind fields – are said not to fit the bill, since what controls action here are propositional – conceptual – *thoughts*, not the kinds of non-conceptual, analogue, representations characteristic of perceptual experience. What Tye overlooks, however, is the way in which perceptual information in the blind field can be involved in detailed, fine-grained, control of movement, such as reaching out to grasp an object (Marcel, forthcoming – but note that these results have in fact been circulating informally for many years). This looks much more characteristic of genuine perception.

⁸ This point is developed at length in my 1992a.

⁹ Another variant of this approach would be to claim that the experience is phenomenally conscious, but is instantaneously forgotten (Dennett 1991). This variant faces problems of its own (see my 1996, ch. 5); and it certainly cannot account for all cases.

¹⁰ In fact I here focus entirely on the question of non-relational definition. For the remaining points, see my 1996, ch. 7. Note that the term ‘qualia’ is sometimes used more neutrally than I do here, as just another way of referring to the *what-it-is-likeness* of experience.

¹¹ What *sort* of sufficient condition? Not conceptual, surely, since the conceivability of zombies suggests that it is conceptually possible for a creature to have all the right representations of its own experiential states while lacking phenomenal consciousness. But to demand a conceptual sufficiency-condition is to place the demands on a naturalistic theory of consciousness too high. We just need to be told what phenomenal consciousness *is*. And a condition which is naturally or metaphysically sufficient can, arguably, do that. See my 1997 for more on the *kind* of account which a theory of phenomenal consciousness has to provide, in order to be successful.

¹² This is what makes me – in one good sense – a *qualia-irrealist*, since I claim that there *are no* non-relational properties of experience *qua* experience.

¹³ Gennaro alleges – surely wrongly – that there is no real distinction between HOE and HOT accounts. In fact the difference supervenes on the distinction between non-conceptual and conceptual content.

¹⁴ Does this have implications for our moral treatment of animals? I once used to think so – see my 1992b, ch. 8. But I now no longer do – see my *forthcoming*. My present view is that it is *first-order* (non-phenomenal) disappointments and frustrations of desire which are the most basic objects of sympathy and (possible) moral concern. (I still think that it is a distinctively moral question – to be addressed by moral theory – whether we are *required* to extend moral concern to animals; and on this my views have not changed.)

¹⁵ There is at least this much truth in so-called ‘simulationist’ accounts of mental-state attribution. See many of the papers in Carruthers and Smith (eds.) 1996.

¹⁶ Many developmental psychologists would say that under the age of *four* most children lack a concept of false belief, and the related concepts of *seeming*, of *subjectivity*, and of *appearances*. I make the claim more cautiously, because increasingly sophisticated experimental techniques continue to push the age of ‘theory-of-mind’ acquisition lower; and because there is evidence that many younger children at least have an *implicit* conception of false belief. See Clements and Perner 1994.

¹⁷ For more detailed development of some of the points made here, see my 1996, chs. 5 and 6.

¹⁸ Dennett (1991) adopted, at the same time, a form of dispositionalist HOR-theory, maintaining that it is a content's *availability* to higher-order thought and description which constitutes it as conscious – and it is because the counter-factuals embedded in the notion of *availability* are thought to lack determinate truth-values that we get the thesis of the radical *indeterminacy* of consciousness. There is an irony here if, as I suppose, the richness of conscious experience provides the main motive for preferring a dispositionalist HOT-theory to its actualist counterpart. It is perhaps no accident, then, that Dennett has now shifted to a form of actualism, saying that consciousness is like *fame* (1995) – constituted by the (actual) effects of a content on surrounding systems, including linguistic description and long-term memory. (And by the way, the indeterminacy of fame is just the indeterminacy of vagueness – there is nothing very radical here any longer.)

¹⁹ Ironically, Dretske himself provides us with just the materials needed, in the chapter (ch. 2) devoted to Introspection. For 'introspection' just read 'potential introspection' throughout, and then you more-or-less have the correct account of phenomenal consciousness!

²⁰ Thanks to David Bell for the initial suggestion for this paper; and to Colin Allen, George Botterill, Susan Granger, Christopher Hookway, Mark Sacks and Michael Tye for comments on earlier drafts. This article is a target-paper for a special volume of the on-line journal *Psyche*, with refereed responses, together with the author's replies, to be published early in 1999 at <http://psyche.cs.monash.edu.au>.

REFERENCES

- Armstrong, D. (1968), *A Materialist Theory of the Mind*. London: Routledge.
- Armstrong, D. (1984), 'Consciousness and causality', in D. Armstrong and N. Malcolm, *Consciousness and Causality*. Oxford: Blackwell.
- Baron-Cohen, S. (1995), *Mindblindness*. Cambridge, MA: MIT Press.
- Block, N. (1995), 'A confusion about a function of consciousness', *Behavioural and Brain Sciences*, 18, 227-247.
- Botterill, G. and Carruthers, P. (1999), *Philosophy of Psychology*. Cambridge: Cambridge University Press.
- Carruthers, P. (1992a), 'Consciousness and concepts'. *Aristotelian Society Proceedings*, Supp. Vol. 66, 41-59.
- Carruthers, P. (1992b), *The Animals Issue: moral theory in practice*. Cambridge: Cambridge University Press.
- Carruthers, P. (1996), *Language, Thought and Consciousness: an essay in philosophical psychology*. Cambridge: Cambridge University Press.
- Carruthers, P. (1997), 'Fragmentary versus reflexive consciousness', *Mind and Language*, 12, 180-194.
- Carruthers, P. (forthcoming), 'Sympathy and subjectivity'.
- Carruthers, P. and Smith, P.K., eds. (1996), *Theories of Theories of Mind*. Cambridge: Cambridge University Press.
- Clements, W. and Perner, J. (1994), 'Implicit understanding of belief', *Cognitive Development*, 9, 377-397.
- Crick, F. and Koch, C. (1990), 'Towards a neurobiological theory of consciousness', *Seminars in the Neurosciences*, 2, 263-275.
- Chalmers, D. (1996), *The Conscious Mind: towards a fundamental theory*. Oxford: Oxford University Press.

- Davidson, D. (1987), 'Knowing one's own mind', *Proceedings and Addresses of the American Philosophical Association*, 60, 441–458.
- Dennett, D. (1978), 'Towards a cognitive theory of consciousness', in his *Brainstorms*, 149–173. Hassocks: Harvester Press.
- Dennett, D. (1991), *Consciousness Explained*. London: Penguin Press.
- Dennett, D. (1995), 'Consciousness: more like fame than television'. Paper delivered at a Munich conference. Published in German as: 'Bewusstsein hat mehr mit Ruhm als mit Fernsehen zu tun', in C. Maar, E. Pöppel, and T. Christaller (eds.), *Die Technik auf dem Weg zur Seele*. Munich: Rowohlt, 1996.
- Dretske, F. (1993), 'Conscious experience', *Mind*, 102, 263–283.
- Dretske, F. (1995), *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Flanagan, O. (1992), *Consciousness Reconsidered*. Cambridge, MA: MIT Press.
- Gennaro, R. (1996), *Consciousness and Self-Consciousness*. Amsterdam: John Benjamins Publishing.
- Harman, G. (1990), 'The intrinsic quality of experience', *Philosophical Perspectives* 4, ed. Tomberlin.
- Humphrey, N. (1986), *The Inner Eye*. London: Faber and Faber.
- Kirk, R. (1994), *Raw Feeling*. Oxford: Clarendon Press.
- Kripke, S. (1972), 'Naming and necessity', in D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, 253–355. Dordrecht: Reidel.
- Lycan, W. (1987), *Consciousness*. Cambridge, MA: MIT Press.
- Lycan, W. (1996), *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Marcel, A. (forthcoming), 'Blindsight and shape perception: deficit of visual consciousness or of visual function?', *Brain*.
- McCulloch, G. (1988), 'What it is like', *Philosophical Quarterly* 38.
- McCulloch, G. (1993) 'The very idea of the phenomenological', *Aristotelian Society Proceedings*, 93.
- McGinn, C. (1991), *The Problem of Consciousness*. Oxford: Blackwell.
- Nagel, T. (1974), 'What is it like to be a bat? ', *Philosophical Review*, 82, 435–456.
- Nagel, T. (1986), *The View from Nowhere*. Oxford: Oxford University Press.
- Nelkin, N. (1996), *Consciousness and the Origins of Thought*. Cambridge: Cambridge University Press.
- Povinelli, D. (1996), 'Chimpanzee theory of mind?', in P. Carruthers and P. K. Smith (eds.), 1996, 293–329.
- Rosenthal, D. (1986), 'Two concepts of consciousness', *Philosophical Studies*, 49, 329–359.
- Rosenthal, D. (1993), 'Thinking that one thinks', in M. Davies and G. Humphreys (eds.), *Consciousness*, 197–223. Oxford: Blackwell.
- Tye, M. (1995), *Ten Problems of Consciousness: a representational theory of the phenomenal mind*. Cambridge, MA: MIT Press.