

Reducing Consciousness by Making it Hot

A Review of Peter Carruthers' *Phenomenal Consciousness*

Robert W. Lurz
Department of Philosophy
Brooklyn College, CUNY
2900 Bedford Avenue
Brooklyn, NY 11210
U.S.A.

rlurz@brooklyn.cuny.edu

Copyright (c) Robert W. Lurz 2002

PSYCHE, 8(05), September 2002
<http://psyche.cs.monash.edu.au/v8/psyche-8-05-lurz.html>

KEYWORDS: Phenomenal consciousness, worldly-subjectivity, experiential-subjectivity, higher-order and first-order theories.

REVIEW OF: Peter Carruthers. (2000). *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge University Press. \$60 hbk. 347 pp. ISBN: 0521781736.

ABSTRACT: Our conscious experiences are said to possess a unique property called *phenomenal consciousness*. Why these and only these states of us have this property has proved to be an exceedingly difficult question for philosophers and scientists to answer. In fact, some have claimed that this question constitutes *the* hard problem of the mind-body problem, one which cannot be solved by the standard methods of contemporary science. In his most recent book, *Phenomenal Consciousness*, Peter Carruthers offers a bold, original and scientifically acceptable solution to this hard problem: the dispositional higher-order thought (HOT) theory. I describe the main line of argument in *Phenomenal Consciousness* for Carruthers' dispositional HOT theory and present three places where the argument seems most vulnerable. I end the review with a very positive endorsement of *Phenomenal Consciousness*, recommending it as compulsory reading for anyone interested in the contemporary philosophical and scientific debate over the nature of phenomenal consciousness.

1. Introduction

Our conscious experiences are said to possess a unique property called *phenomenal consciousness*. Why these and only these states of us have this property has proved to be an exceedingly difficult question for philosophers and scientists to answer. In fact, some have claimed that this question constitutes *the* hard problem of the mind-body problem, one which cannot be solved by the standard methods of contemporary science. Peter Carruthers strongly disagrees. In his most recent book, *Phenomenal Consciousness*, Carruthers offers a bold, original and scientifically acceptable solution to this hard problem. Carruthers' book provides a clear overview and analysis of all the important philosophical theories of phenomenal consciousness on the market, along with many lucid accounts of relevant, up-to-date findings on consciousness from neuroscience and psychology. I highly recommend *Phenomenal Consciousness* to anyone interested in the current philosophical and scientific debate over the nature of phenomenal consciousness.

Phenomenal Consciousness is quite large in scope, covering the full range of contemporary philosophical and scientific views on consciousness -- from Thomas Nagel's 1974 paper, "What is it like to be a bat?", to recent work by Lawrence Weiskrantz on blindsight. I do not intend, therefore, to give a full account of all that takes place in the book. Rather, I would like to give a rough overview of (what I take to be) the book's main line of argument, leaving my few criticisms for the very end.

2. Overview of *Phenomenal Consciousness*

In *Phenomenal Consciousness*, Carruthers puts forward and defends a dispositional higher-order thought (HOT) theory of phenomenal consciousness. Roughly, the theory states that "phenomenal consciousness consists in a certain sort of intentional content ('analog', or fine-grained), held in a special-purpose short-term memory store in such a way to be available to higher-order thoughts about the occurrence and nature of those contents; and that in virtue of such availability (given the truth of some or other form of 'consumer semantics') all of those contents are at the same time higher-order ones, acquiring a dimension of *seeming* or *subjectivity*" (p. xiii).

Carruthers takes his theory to provide a "reductive explanation" of phenomenal consciousness, by which he means that it "describes a way of linking together cognitive structures and contents (in terms which do not themselves presuppose phenomenal consciousness), any instantiation of which is supposed to be metaphysically sufficient for phenomenal consciousness to occur" (p. 257). There are no possible worlds, according to Carruthers, in which the cognitive structures and contents described in his account are appropriately instantiated in a creature's mind and that creature fails to undergo phenomenally conscious states. Carruthers also takes his dispositional HOT account to describe "the actual natural *constitution* of phenomenal consciousness, in this world" (p. 259). In our world, any creature that fails to appropriately instantiate these cognitive

structure and contents does not, according to Carruthers, have phenomenally conscious states.

Carruthers has two main objectives in his book: (1) to defend his account against alleged counterarguments designed to show that the cognitive structures and contents described in his account are either metaphysically insufficient for phenomenal consciousness (arguments for *qualia*) or are not constitutive of phenomenal consciousness in this world (arguments for phenomenal consciousness in animals and infants), and (2) to show that his account is superior to any of the major rival naturalistic accounts of phenomenal consciousness currently on the market. Before an overview is given of these objectives, a few words need to be said about the intended *explanandum* of Carruthers' account.

2.1. On The Definition of "Phenomenal Consciousness"

Phenomenal consciousness, according to Carruthers, "is the property which mental states possess when it is *like something* to have them" (p. 13). To illustrate, imagine that you are looking at a red tomato and a green lime under good lighting conditions. There is something that it is like for you to see the red tomato, and what it is like is different from what it is like for you to see the green lime. There is a subjective difference between your seeing the tomato and your seeing the lime, and this difference is a difference in phenomenal consciousness.

Things are not as clear as this illustration might suggest, however. For Carruthers makes a rather subtle distinction between two senses of the phrase 'what it is like', only one of which is to be identified with phenomenal consciousness. According to Carruthers, there is a conceptual and a real distinction between "what the *world* (or that 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" (p. 128). The former sense of subjectivity Carruthers calls "worldly-subjectivity"; the latter sense he calls "mental-state-subjectivity" or "experiential subjectivity". It is this latter sense of subjectivity that Carruthers identifies as phenomenal consciousness (p. 129, n. 7).

Upon first encountering this rather subtle but important distinction between worldly and experiential subjectivity, the reader may struggle to make sense of it -- at least, this reader did. And I found that the distinction was given its clearest expression in Chapter 6, where Carruthers puts forward an argument -- drawing on cases from everyday life and scientific research -- for the existence of *non-conscious* experiences. The upshot of the chapter -- at least, with respect to its bearing on the worldly-subjectivity/experiential-subjectivity distinction -- is that non-conscious experiences have worldly-subjectivity without having experiential-subjectivity; whereas, conscious experiences have both types of subjectivity. To illustrate the conscious/non-conscious distinction and its bearing on the worldly-subjectivity/experiential subjectivity distinction, let me appeal to two rather well-known cases which Carruthers uses to show the existence of non-conscious experiences: the absent-minded driver case and the blindsight case.

The Absent-Minded Driver Case: An absent-minded driver is paying little or no attention to what he is seeing on the road, and yet he keeps his car on the road and successfully navigates it around other vehicles. Imagining himself as such a driver, Carruthers maintains that:

I surely must have been seeing, or I should have crashed the car. Indeed, my passenger sitting next to me may correctly report that I *saw* the vehicle double-parked at the side of the road, since I deftly turned the wheel to avoid it. Yet I was not conscious of seeing it, either at the time or later in memory. My percept of that vehicle was not a conscious one. (p. 149)

The Blindsight Case: In cases of blindsight, subjects who have areas of their striate cortex damaged claim to see nothing in certain blind areas (scotomas) of their visual fields but retain the remarkable ability (manifested on forced-choice test trials) to visually discriminate objects and features within their scotomas. Blindsight subjects can, Carruthers reports, "accurately trace the movement of a light across a screen in the blind portion of their visual field, entirely by guess-work, professing that they are aware of nothing; and they can also discriminate colours without conscious awareness" (p. 155).

In both these cases, it seems, the subjects are (in some sense) visually aware of items in their environment -- the absent-minded driver is aware of the double-parked vehicle, and the blindsight subject is aware of the movement of the light across the screen. The world in both these cases, according to Carruthers, is "*subjectively presented*" (via the stimulation of the subjects' relevant sensory organs) to the subjects as being a certain way. In this respect, Carruthers maintains, their visual experiences have worldly-subjectivity (p. 128). However, as Carruthers points out, these subjects are not *consciously* aware of these items in their environment: their respective perceptual experiences occur below the level of consciousness. Consequently, it is quite intuitive to say, as Carruthers does, that there is nothing that it is like for these subjects to undergo these non-conscious perceptual experiences (p. 176). Their perceptual experiences, then, lack what Carruthers calls experiential subjectivity or phenomenal consciousness.

With the worldly-subjectivity/experiential-subjectivity distinction sketched, let us return to your visual experiences of the red tomato and the green lime. Your visual experiences, we shall assume, are conscious. You are not suffering from absent-mindedness or blindsightedness, and you are focusing your attention on the tomato and the lime before you. Now, your visual experience of the tomato represents the tomato as being a certain way (namely, as being red and round) and, as a result, has worldly-subjectivity. But, there is also something that it is like for you to undergo this visual experience, and, as a result, it also has experiential-subjectivity (or phenomenal consciousness). The same can be said, *mutatis mutandis*, for your visual experience of the lime.

However, the experiential subjectivity of your visual experience of the red tomato is distinct from that of your visual experience of the green lime: what it is like for you to see the red tomato is different from what it is like for you to see the green lime. An important part of Carruthers' project in *Phenomenal Consciousness* is to give a naturalistic account

of this type of distinction by appealing only to cognitive structures and intentional contents of mental states that do not presuppose the existence of phenomenal consciousness. And, as indicated above, Carruthers attempts to do this by showing that the intentional contents of your conscious experiences are different. The difference in the experiential subjectivity between your consciously seeing the red tomato and your consciously seeing the green lime resides entirely in the difference between the contents of your experiences. And this difference in the intentional contents of your experiences resides, in part, in the fact that your experiences -- since they are conscious -- are available to your higher-order thought (HOT) module in your brain.

2.2. Some Important Details of Carruthers' Account

Some important details of Carruthers' theory need to be added at this point; and we can do this, in part, by looking at the following diagram which Carruthers provides.

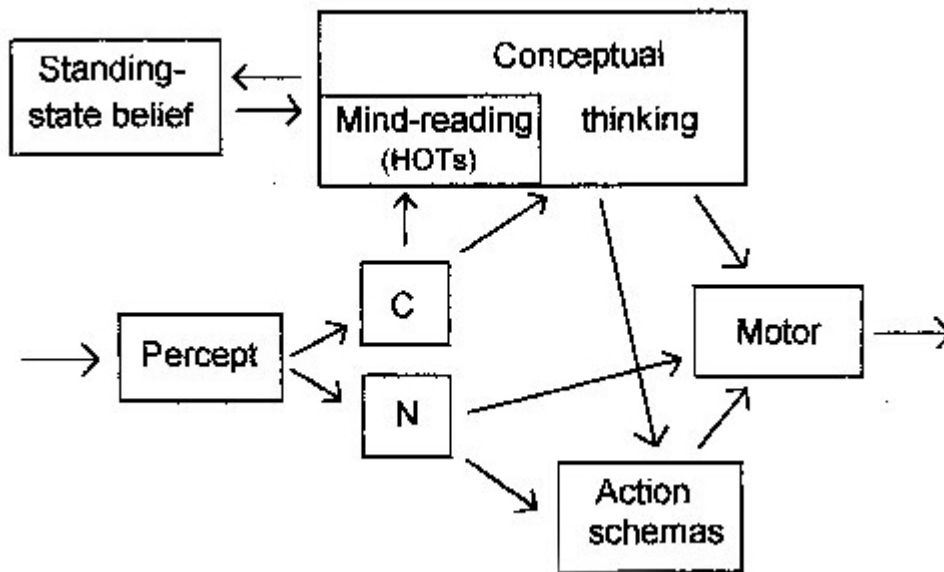


Figure 1.

Dispositional HOT Theory

According to Carruthers' theory, conscious experiences -- such as your conscious visual experience of the tomato -- are percepts (or analog/fine-grained representations) that represent first-order features (redness, roundness, spatial relations, etc.) of objects in the environment. And conscious experiences (percepts) are made conscious in virtue of their being made available to the subject's higher-order thought (HOT) forming module via a special short-term memory store C. Non-conscious experiences -- such as those of the absent-minded driver and blindsighted subject -- are not made available to the subject's HOT forming module by way of being held in C; rather, they are held in a different short-

term memory store N, from which they are made available to the subject's motor-output system, either directly or by way of the subject's action-schemas module.

On Carruthers' account, both conscious and non-conscious percepts may represent the very same first-order features of external objects. A blindsighted subject's non-conscious visual experience of a red tomato, for example, may represent the tomato as being red and round, just as your conscious visual experience of the tomato represents the tomato as being red and round. This commonality in the first-order contents of these experiences explains, according to Carruthers, the fact that what the *world* is like for the blindsight subject (whilst she is aware of the tomato in her blind region) is (or may be) what the *world* is like for you (whilst you are aware of the tomato). Furthermore, since your conscious visual experience of a green lime represents a different first-order property (namely, greenness) from that represented (we are supposing) in the blindsighted subject's non-conscious visual experience of the red tomato, what the *world* is like for you (whilst you are aware of the lime) is not what the *world* is like for the blindsighted subject (whilst she is aware of the tomato in her blind region). First-order contents of percepts, on Carruthers' account, are taken to explain the *worldly-subjectivity* of experiences.

But, as we saw above, there is a difference in the *experiential subjectivity* between conscious and non-conscious experiences, and a difference in the *experiential subjectivity* among conscious experiences of different types (e.g., the difference between consciously seeing a red tomato and consciously seeing a green lime), which needs to be explained. Carruthers' explanation of these differences runs roughly as follows. The fact that a conscious experience is available to a subject's HOT module -- that is, available to the subject to think about it -- adds a higher-order dimension to the experience's intentional content. Conscious experiences, on Carruthers' account, not only represent first-order properties of objects in the environment (as non-conscious experiences do), but they also represent higher-order properties -- properties of the experience itself. Carruthers explains:

Where before these were first-order analog representations of the environment (and body), following the attachment of a HOT system these events take on an enriched dual content. Each experience of the world-body becomes at the same time a representation that just such an experience is taking place; each experience with the content *red(a)*, say, is at the same time an event with the content *seems red(a)* or *experience of red(a)*.[<1>](#) (p. 242)

The difference in the experiential subjectivity between *conscious* and *non-conscious* experiences, according to Carruthers, resides entirely in the difference in the intentional contents of these types of experiences. Non-conscious experiences have only first-order contents; whereas, conscious experiences have first-order and higher-order contents. And the difference in experiential subjectivity among *conscious* experiences of different types resides entirely, according to Carruthers, in the different higher-order contents of these experiences:

The difference between the *feel* of a phenomenally conscious experience with the analog content *red(a)*, on the one hand, and the *feel* of a phenomenally conscious experience with the analog content *green(a)*, on the other, is that the former also has the higher-order content *seems red(a)* whereas the latter also has the higher-order content *seems green(a)*. So each phenomenally conscious experience has its distinctive form of subjectivity by virtue of acquiring a higher-order analog content which precisely mirrors, and represents *as* subjective, its first-order content. (p. 243)

Finally, it should be noted that Carruthers takes the intentional contents of percepts to be individuated narrowly -- in abstraction from the objects and properties in a creature's environment -- and amenable to a naturalistic explanation in terms of the narrow functional roles that these percepts play within the mind of the creature that possesses them -- what Carruthers calls "consumer semantics" (pp. 95-112). The narrow functional role of *conscious* percepts -- specifically, their availability to a creature's HOT module via a special short-term memory store C -- metaphysically determines their unique, narrow, higher-order intentional contents, which, in turn, metaphysically determines their possession of phenomenal consciousness. On Carruthers' account, phenomenal consciousness boils down to a property that percepts possess in virtue of a unique, narrow functional role that they play within the mind of the creature that possesses them.

2.3. Two Main Tasks in *Phenomenal Consciousness*

Carruthers sets himself two main tasks in his book, as mentioned above. The first is to defend his account against alleged counterarguments, and the second is to show that his account is superior to any other major rival account of phenomenal consciousness on the market. With regard to the first task, Carruthers addresses (among others) four traditional counterarguments against higher-order/functionalists' accounts of phenomenal consciousness: the absent-*qualia* argument, the inverted-*qualia* argument, the inverted-earth argument, and the argument for phenomenal consciousness in animals and infants.

The first two arguments purport to show that it is metaphysically possible -- by way of its being conceivable -- that there be a creature A who is functionally isomorphic to a creature B, who realizes the functional architecture in Carruthers' account (as illustrated in Figure 1 above), and yet A's percepts either lack phenomenal consciousness altogether (the absent-*qualia* argument) or the phenomenal character of A's percepts are inverted relative to B's (the inverted-*qualia* argument). The third counterargument (the inverted-earth argument) purports to show that it is metaphysically possible -- again, by way of its being conceivable -- that there be a creature A who is functionally heteromorphic to a creature B, who realizes the functional architecture in Carruthers' account, and yet A's percepts have the same phenomenal character relative to B's. Were either of these three alleged metaphysical possibilities genuine, they would show that the functional

architecture described in Carruthers' account fails to metaphysically fix the nature and existence of phenomenal consciousness.

Carruthers responds to the absent- and inverted-*qualia* arguments by showing that at most these arguments prove the *conceptual* possibility of such cases of absent and inverted *qualia*; but they do not prove the *metaphysical* possibility of such cases, which is what is needed to disprove his account. In his response to these two arguments, Carruthers relies upon a sharp distinction between the *property* of phenomenal consciousness (the *explanandum* of his account) and the *concept* of phenomenal consciousness (which he takes to be a recognitional concept). At most, the conceivability of such cases described in these two counterarguments, Carruthers maintains, shows that our recognitional concept of phenomenal consciousness cannot be functionally defined; they do not show that the *property* that this concept picks out cannot be a functional property. <2>

With regard to the inverted-earth argument, Carruthers maintains that such cases -- cases in which an earthling is given inverted-color lenses and brought to a place where the colors of things are inverted relative to their counterparts on earth -- merely shows that the *wide* causal/functional roles of percepts do not metaphysically fix their phenomenal character; they do not show that the *narrow* causal/functional roles do not. And it is the latter metaphysical possibility that must be shown, Carruthers maintains, in order to disprove his account.

Finally, the fourth counterargument, to which Carruthers considers and responds, purports to show that the higher-order architecture described in his account is not constitutive of phenomenal consciousness in our world. According to this counterargument, it is intuitively plausible that many animals and young children possess phenomenally conscious experiences. That is, it is intuitively plausible that there is something that it is like for (say) a cat or a two-year-old to see a scurrying mouse or a red ball, and that what it is like for each is different from what it is like for each to see a stationary mouse or a green block, or to hear a trumpet blast, or to feel an itch, and so on. However, it is also plausible -- and there is some rather suggestive scientific evidence (see Clements and Perner (1994) and Povinelli (1996)) that corroborates this -- that these creatures do *not* have the ability to form higher-order thoughts about their own perceptual states and, hence, do not have minds that contain the HOT module which, on Carruthers' account, is essential for phenomenal consciousness.

Carruthers' responds to this counterargument by attempting to show that the intuitive belief behind it is groundless and easily explained away. <3> The intuition that animals and young children possess phenomenally conscious experiences, according to Carruthers, is an illusion, resulting from the way we attribute mental states to these creatures. Using the attribution of an experience to a cat as a case in point, Carruthers explains the mechanism behind this illusion as follows:

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 that experience might be like 'from the inside'... 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 then bring with it its distinctive subjective phenomenology. So we are subject to a kind of cognitive illusion -- an illusion which arises because we cannot consciously imagine a mental state which is *unconscious* and lacking any phenomenology. When we imagine the mental states of non-human animals we are necessarily led to imagine states which are phenomenological; and this leads us to assert (falsely, if the arguments given here are correct) that if non-human animals have any mental states at all (as they surely do), then their mental states must be phenomenally conscious ones. (p. 198)

Now, there are theories of phenomenal consciousness on the market -- specifically, first-order (FO) theories -- that are consistent with and offer an explanation of the intuition that animals and young children have phenomenally conscious percepts. Carruthers, however, attempts to show that such theories suffer serious problems which his own dispositional HOT account does not. And this brings us to the second main task in Carruthers' book: showing that his account is superior to any other major rival account of phenomenal consciousness on the market. Carruthers describes and objects to no less than six different contemporary accounts of phenomenal consciousness: three different kinds of FO accounts (which he attributes to Fred Dretske, Michael Tye, and Robert Kirk, respectively) and three different kinds of higher-order (HO) accounts (which he attributes to David Armstrong, David Rosenthal, and Dan Dennett, respectively). Much is to be learned from Carruthers' penetrating analysis and criticisms of these rival accounts, but for the sake of brevity, I shall focus on his analysis and criticism of FO accounts.

Roughly, FO accounts attempt to explain the nature of phenomenal consciousness in terms of an experience's first-order intentional content and first-order functional role. The main goal of FO accounts, according to Carruthers, is to

Characterise all of the phenomenal -- 'felt' -- properties of experience in terms of the representational contents of experience (widely individuated). On this view, the difference between an experience of red and an experience of green will be explained as a difference in the properties represented -- reflective properties of surfaces, say -- in each. (p. 114)

On FO accounts, the phenomenal difference (for instance) between your conscious visual experience of a red tomato and your conscious visual experience of a green lime resides entirely in the different first-order contents of these experiences. Your visual experience of the tomato represents an external object as having the properties of being round and red; whereas, your visual experience of the green lime represents an external object as being round and green. Although FO theorists disagree over what fixes an experience's first-order representational content, they agree that there is no need to attribute (as Carruthers does) higher-order representational contents to experiences -- contents about the experiences themselves -- in order to account for the phenomenal difference among different experiences.

FO theorists acknowledge, however, that there are some mental states which possess intentional contents, such as beliefs and thoughts, but do not possess phenomenal properties. According to FO theorists, the difference between these mental states and sensory states, which do possess phenomenal properties, resides in the distinct first-order functional role that sensory states play in the mind. According to FO theorists, sensory states, unlike beliefs and thoughts, are (to use Carruthers' words) "the *output[s]* of the various (peripheral) perceptual systems (sight, hearing, touch, etc.), which [are] presented as *input[s]* to the various (central) cognitive systems charged with fixing [first-order] beliefs, generating plans, and controlling movements" (p. 115). Carruthers captures the essentials of the functional architecture of the FO account with the following diagram:

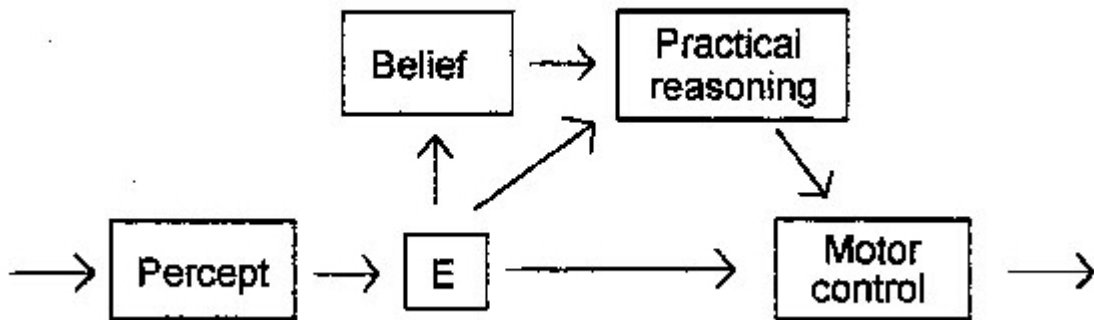


Figure 2.

First-Order Representationalism

The sensory states that are contained in a subject's E box, according to FO theorists, are phenomenally conscious states in virtue of their availability to the subject's first-order belief-forming and practical-reasoning modules. And the different sensory states that reside in a subject's E box differ in phenomenal character, according to FO theorists, in virtue of their distinct first-order intentional contents.

The chief problem for FO accounts, according to Carruthers, is their inability to give an adequate account of the conscious/non-conscious distinction for experiences. Recall that, for Carruthers, phenomenal consciousness is possessed by, and only by, *conscious* experiences. FO theorists, Carruthers argues, are faced with an intractable trilemma (pp. 168-179). They can either (i) accept that all and only conscious experiences are phenomenally conscious but deny that there is any genuine conscious/non-conscious distinction for experiences to be explained; (ii) accept that all and only conscious experiences are phenomenally conscious, accept that there is a genuine conscious/non-conscious distinction for experiences to be explained, and attempt to explain this distinction in terms of a first-order functional architecture; or (iii) deny that all and only conscious experiences are phenomenally conscious, accept that there is a genuine conscious/non-conscious distinction for experiences to be explained, and attempt to explain this distinction in terms of a higher-order functional architecture. <4>

Neither of these three options, Carruthers argues, is plausible. Option (i) runs afoul of certain facts --- gleaned from everyday cases and scientific studies -- which strongly suggests the existence of non-conscious experiences (pp. 147-168); option (ii) fails to explain why there should be anything that it is like for a subject to possess a conscious (as opposed to a non-conscious) experience (pp. 170-174); and option (iii) is committed to the highly counterintuitive view that non-conscious experiences are phenomenally conscious (pp. 175-179). The only plausible way out of this trilemma, Carruthers suggests, is to abandon FO accounts entirely and accept a HO account of phenomenal consciousness, such as his own dispositional HOT account.

3. Problems for *Phenomenal Consciousness*

Notwithstanding its many virtues, the main chain of reasoning in *Phenomenal Consciousness*, I believe, has a few weak links. In this section, I wish to point out three of these links.

3.1. Taking the Trilemma by a Horn

Insofar as I understand Carruthers' trilemma against FO accounts, I do not see how the FO theorist could escape from it by passing through its horns. Rather, it seems to me that to escape from the trilemma, the FO theorist must somehow blunt one of its horns; and of the three horns presented, option (ii) seems the most intuitively plausible. So, it seems to me that the FO theorist should, if he can, take up and defend option (ii). But, contrary to what Carruthers maintains, I believe that the FO theorist can do this without impaling himself on option (ii)'s point.

Carruthers maintains that any FO theorist who takes up option (ii) fails to explain why conscious and only conscious experiences are phenomenally conscious. It is not entirely clear to me, however, why Carruthers thinks that this is so. But, on one plausible reading, he appears to reason as follows. Any FO theorist who takes up and defends option (ii) must give an account of the conscious/non-conscious distinction for experiences which does not appeal to higher-order mental states, such as the state of being aware of one's own experiences. On such an account, then, conscious experiences need not be experiences that their possessors are (or are able to be) aware of. But, this, then, seems to make it mysterious why conscious experiences should be any more phenomenally conscious than the non-conscious experiences of blindsight subjects and absent-minded drivers.<5> For blindsight subjects and absent-minded drivers, recall, are not (and are unable to be) aware of their respective experiences, and it is precisely this fact about these individuals which makes it intuitively plausible to say that their respective experiences are not phenomenally conscious (p. 176). Phenomenally conscious experiences -- experiences that there is something-it-is-like to have -- are, intuitively, experiences that their possessors are (or are able to be) aware of (p. 14). Therefore,

without an appeal to higher-order mental states -- specifically, the state of being aware of one's own experiences -- the FO theorist who takes up option (ii) seems to be unable to explain why conscious experiences, and only conscious experiences, should be phenomenally conscious.

If this is Carruthers' reasoning on pages 170-174, then I find it compelling *only if* there is but one reading of the expression 'S is aware of his experience E' which makes the claim 'phenomenally conscious experiences are experiences that their possessors are (or are able to be) aware of' true, and that that reading is a higher-order reading. On a higher-order reading, to say that S is aware of E is (roughly) to say that S is aware of E *as such* -- that is, that S is aware that he has an *experience E*. But, is this the only plausible reading of 'S is aware of E' which makes the claim 'phenomenally conscious experiences are experiences that their possessors are (or are able to be) aware of' true? I'm not sure that it is. In fact, there appears to be a non-higher-order reading of this expression which is just as likely to make this claim true as the higher-order reading. On this non-higher-order reading, to say that S is aware of E is simply to say that S is aware *of what* E represents.<6> And to be aware *of what* something, *x*, represents is not the same as being higher-order aware that *x* represents such-and-such.

To illustrate this difference, consider an analogous distinction in cases of *trompe l'oeil* paintings. When one is fooled by a *trompe l'oeil* painting -- say, a painting of an outdoor scene -- one is aware of the painting but not *as such*. One is not aware that there is a *painting* of an outdoor scene present. For if one were aware of the painting in this way, one would not be fooled by it. Rather, when one is fooled by the *trompe l'oeil* painting, one is aware of the painting simply in the sense of being aware *of what* it represents. In such cases, one is aware *of what* is represented in the painting -- an outdoor scene, for example -- without being aware that a *painting* represents it.

Analogously, being aware *of what* one is experiencing (when one is experiencing an *F*) is not the same as being aware that one is *experiencing* an *F* (when one is experiencing an *F*). The latter is a higher-order mental state -- one that represents a lower-order mental state (one's experience of an *F*) *as such* -- but the former is not. Nevertheless, in being in the former mental state -- in being aware *of what* one is experiencing (when one is experiencing an *F*) -- one is, in one sense, aware of one's experience. One is aware of one's experience in just the same way as one is aware of the *trompe l'oeil* painting when one is aware *of what* it represents. Normal sighted subjects are aware of many of their conscious visual experiences in this non-higher-order sense of 'aware of'. When I consciously see a red tomato, for instance, I am aware *of what* (N.B.: not *that*) I am seeing, as evidenced by the fact that I can sincerely say, describe, or remember, *what* (N.B.: not *that*) I am (was) seeing.<7> And in the non-higher-order sense of 'aware of', I am thereby aware of seeing a red tomato.

However, not all subjects who have visual experiences of items or features are (or are able to be) aware *of what* they are experiencing when they are experiencing these items or features. Blindsight subjects, recall, are able to experience some features within their blind regions, as evidence by their performance on forced-choice test trials. But, these

subjects are not aware *of what* (N.B.: not *that*) their visual experiences represent, as evidenced by their inability to sincerely say, describe, or remember *what* (N.B.: not *that*) they are (were) experiencing.<8> These subjects, in a word, are unaware *of what* they are in fact visually aware of. In the non-higher-order sense of 'aware of,' then, blindsight subjects are *not* aware of seeing those features within their blind regions. A similar description can be given, *mutatis mutandis*, for the non-conscious visual experiences of absent-minded drivers.<9>

I see no reason why the FO theorist cannot use this non-higher-order reading of 'S is aware of E' as a way of explaining why, on his FO account, all and only conscious experiences are phenomenally conscious. The non-conscious visual experiences of blindsight subjects and absent-minded drivers are not phenomenally conscious, according to the FO theorist, because the possessors of these experiences are wholly unaware *of what* these experiences represent; whereas, the conscious visual experiences of normal subjects are phenomenally conscious, according to the FO theorist, because the possessors of these experience are (or are able to be) aware *of what* these experiences represent. Put bluntly: there is something that it is like to have an experience E if one is (or is able to be) aware *of what* E represents, but there is nothing it is like to have an experience E if one is not (or is unable to be) aware *of what* E represents.

The mystery for FO accounts of why conscious experiences should be any more phenomenally conscious than the non-conscious experiences of blindsight subjects and absent-minded drivers seems to disappear once this non-higher-order reading of 'S is aware of E' is made explicit and pressed into service. It is true, of course, that on such an account, conscious experiences need not be experiences that their possessors are (or are able to be) aware of, in the higher-order sense of 'aware of'. But, this leaves it mysterious why conscious experiences should be any more phenomenally conscious than the non-conscious experiences of blindsight subjects and absent-minded drivers *only if* one assumes that a subject must be (or must be capable of being) higher-order aware of E in order for E to be phenomenally conscious. But, such an assumption merely begs the question against a FO account of phenomenal consciousness. And since there appears to be a legitimate non-higher-order reading of 'S is aware of E' which seems to explain why all and only conscious experiences are phenomenally consciousness, Carruthers must show, and not merely assume, that on such a reading, FO accounts still fail to explain why all and only conscious experiences are phenomenally conscious. On the face of it, the FO theorist seems equipped to blunt the second horn of Carruthers' trilemma.<10>

3.2. Explaining Away an Intuitive Belief

I'm not sure about other people, but I believe that many animals have conscious perceptual experiences and beliefs about items and facts in their environments. My reason for holding this belief is not that I have succeeded in imagining what it is like for animals to perceive or have beliefs about items and facts in their environments - for I have not. Rather, my reason is simply that a lot of animal behavior can be predicted and

explained quite well in terms of the concepts and generalizations of our folk psychology, and these concepts and generalizations are about *conscious* mental states. Our folk theory is a theory of how conscious mental states interact and produce behavior. It is not a theory about how unconscious mental states interact and produce behavior. Unconscious mental states fall under their own laws, and it is the project of scientific psychology to discover them. Therefore, since our folk theory does a fairly good job in explaining and predicting quite a lot of animal behavior, there is some grounds to believe that animals have the kinds of mental states that our folk theory's concepts and generalizations are about - that is, conscious mental states.

Carruthers claims that such a belief is, in fact, groundless, since there is no real need to attribute conscious mental states to animals in order to explain their behavior. He writes:

Everything that the cat does can be explained perfectly well by attributing beliefs, desires, and perceptions to it. There is no explanatory necessity to attribute *conscious* beliefs, desires, or perceptions. 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. (p. 199)

But, to what theory does Carruthers think that these concepts of belief, desire, and perception belong? If they belong to our ordinary folk theory, then they are concepts of conscious beliefs, desires, and perceptions. On the other hand, if they belong to some other theory - say, a scientific theory about unconscious beliefs, desires, and perceptions - then clearly the burden is on Carruthers to demonstrate that everything that animals do can be explained perfectly well with this theory. But as far as I know, no such demonstration has been made. And this is so, in part, because there simply isn't at present a fully developed scientific theory of unconscious beliefs, desires, and perceptions with which to run such a demonstration. But, until someone produces such a scientific theory and demonstrates that it explains everything that animals do, Carruthers, it seems, has given little reason to think that my belief in animal conscious mentality is groundless.

I am also not convinced that Carruthers succeeds in explaining away my belief that many animals have conscious mental states. According to Carruthers, recall, I hold this belief because whenever I attribute such mental states to animals, I undergo -- "almost habitually" -- the following sort of process of imagination:

- i. I imagine "from the inside" what it is like for the animal to possess the mental state which I attribute to it;
- ii. in doing (i), I inevitably imagine what it is for *me* to possess this mental state; and
- iii. in doing (ii), I inevitably imagine a *conscious* mental state.

As a result of doing (iii), I am naturally led, Carruthers maintains, to "assert that if non-human animals have mental states at all (as they surely do), then their mental states must be phenomenally conscious" (p. 198).

What I find puzzling about Carruthers' explanation here is that it fails to describe what typically goes on in my mind when I attribute mental states to animals. Typically, I do not find myself trying to imagine "from the inside" what it is like for an animal to have the mental states that I attribute to it. When I say that the cat sees the mouse, or that the dog smells the buried truffle, I do not usually find myself trying to imagine what it is like for these animals to have their respective perceptual experiences. And in many cases, I know full well that I could not possibly imagine what it is like for an animal to have the mental states that I attribute to it. I am quite sure, for instance, that I could not imagine what it is like for a dog to smell a buried truffle, and so I do not even try to imagine this. Nevertheless, I certainly believe that dogs do smell buried truffles and that their olfactory experiences in such cases are conscious. Carruthers' explanation, then, would appear to be unable to account for why I should find it so natural to attribute conscious mental states to animals when I do not (and, in some cases, can not) imagine "from the inside" what it is like for the animal to have the mental states that I attribute to it.

Carruthers offers the following reply to this puzzling feature of his explanation:

[I]t is plausible that what happens in such cases is that we generalize from those experiences which we *can* imagine. Since any experience which we can imagine must have the properties of feel and what-it-is-likeness, we naturally assume that any experience whatsoever (even those which we cannot imagine) must be similarly phenomenally conscious. (p. 199)

But, if I do engage in such acts of generalization, I do not seem to be aware of them. I do not recall consciously making such a generalization. And it is certainly not the sort of explanation that I would give for why I believe that many animals have conscious mental states. Furthermore, it should be noted that I am not prone to such acts of generalizations in other domains. For instance, I am not prone to think that all closed-sided figures have fewer than eight sides simply because all closed-sided figures that I can imagine have fewer than eight sides, or that all music of Beethoven's is but four notes long simply because all music of Beethoven's that I can imagine is but four notes long, and so on. So, why think it plausible to suppose that I (and others) engage in this sort of generalization when it comes to the mental states of animals? It seems to me that a better explanation for why I (and perhaps others) believe that many animals have conscious mental states is that a lot of animal behavior can be predicted and explained quite well in terms of the concepts and generalizations of our folk psychology, and these concepts and generalizations are about *conscious* mental states. At least, this is the sort of explanation that I would offer for my belief.

3.3. The Transparency Theses

Carruthers' dispositional HOT theory of phenomenal consciousness seems to run afoul of a particular version of the transparency thesis. One version of this thesis (endorsed by Harman (1990), Tye (1995), and Dretske (1995)) states that there are no introspectible, non-representational features of sensory experiences: when we introspect a sensory experience, we are not aware of any feature of our experience over and above those features which our experience represents. Call this the *representational transparency thesis*. There is, however, another, stronger version of the transparency thesis, which we can call the *first-order transparency thesis*. According to this thesis, when we introspect a sensory experience, we are not aware of any feature of our experience over and above those first-order features which our experience represents -- where first-order features are those features possessed by external, non-mental objects.^{<11>} So, for instance, when I see a red tomato and am introspectively aware of seeing a red tomato, I am not aware of any feature of my visual experience over and above the first-order properties it represents the tomato as having -- such as, the particular color, shape, location, texture, etc. of the tomato. I am not, in other words, introspectively aware of having a visual experience as of a *visual experience* of a red tomato; I am simply introspectively aware of having a visual experience as of a *red tomato*.

Carruthers' dispositional HOT account appears to be perfectly consistent with the representational transparency thesis. After all, part of his project in *Phenomenal Consciousness* is to explain away those putative introspectible, non-representational features of experiences (p. 94). But, Carruthers' account appears to be inconsistent with the first-order transparency thesis. For on his account, every conscious experience is not only a first-order representation of features of external objects but a "higher-order experience" which represents "that just such an experience is taking place; each experience with the content *red(a)*, say, is at the same time an event with the content *seems red(a)* or *experience of red(a)*" (pp. 242-3). On Carruthers' account, then, it would appear that when I am introspectively aware of seeing a red tomato, I am not only aware of having a visual experience as of a red tomato, I am also aware of having a visual experience as of a *visual experience* as of a red tomato.

I say that Carruthers' account *appears* to be inconsistent with the first-order transparency thesis because he is not entirely clear on the nature of the analog mode of presentation *seems red(a)* contained in the dual content of conscious experiences. On most occasions, Carruthers takes this analog mode of presentation to be equivalent to the mode of presentation *experience of red(a)*, which is a higher-order mode of presentation -- presenting to the subject a property of the subject's *experience* (namely, the property of being of *red(a)*). On this reading of *seems red(a)*, Carruthers' account is clearly inconsistent with the first-order transparency thesis. But, sometimes Carruthers appears to take *seems red(a)* as presenting to the subject a first-order property of the distal object of the experience. In this way, *seems red(a)* is on par with the first-order analog mode of presentation *red(a)*, in that both present to the subject a property of the distal object of the experience. Carruthers appears to treat *seems red(a)* in this first-order way when he attempts to explain "the so-called 'transparency' of perceptual experience" in terms of the machinery of his HOT account:

This is now easily and satisfyingly explained. The reason why you do not discover any additional properties of your experience when you concentrate your attention on it -- in addition, that is, to the properties of the world (or body) represented -- is that there are no such properties. All that happens when you focus your attention on your experience of the ripe red tomato is that you attend to a state with an analog content representing *redness(a)*, which also represents *seeming redness(a)*. And either way, to focus your attention on this state is to focus on the redness represented. (p. 248-9)

But, if all I am introspectively aware of, when I am introspectively aware of seeing a red tomato (for example), are the properties of the world (in this case, the properties of the tomato) represented in my visual experience, then *seems red(a)* and *red(a)* must present to me properties of the world (the tomato), not properties of my visual experience of the world (the tomato). Hence, in the above paragraph, *seems red(a)* and *red(a)* appear to be taken as first-order modes of presentation of the world (the tomato).

Taking *seems red(a)* in this first-order way certainly allows Carruthers' account to be consistent with the first-order transparency thesis, but it does so only to saddle it with other problems. First, if *seems red(a)* and *red(a)* both present the tomato as having a certain property, the question arises whether these properties are the same or different. If they are the same, then one would like to know how *seems red(a)* and *red(a)* are different? Carruthers, I would think, would want to individuate these different modes of presentation in terms of their different functional roles. But, the problem here is to say how their functional roles would differ on his HOT account. Since *seems red(a)* is not -- we are supposing here -- a higher-order mode of presentation, it would seem unlikely that it should be functionally individuated in terms of its availability to a creature's HOT module. On the other hand, if these modes of presentation pick out different properties, then one would like to know what these two different properties are and how they are related. We have some rough idea of what the property of the tomato is which is picked out by *red(a)*, but what is the property of the tomato which is picked out by *seems red(a)*? And how are these two properties related? Does one supervene on the other, or are they merely correlated?

Finally, and perhaps more importantly, if *seems red(a)* is a first-order mode of presentation, then why think that one needs to appeal to a HOT module in a creature's mind in order to account for its first-order representational content? And if one need not appeal to a HOT module in a creature's mind in order to account for the content of *seems red(a)*, and if the content of this mode of presentation is (in part) what constitutes the phenomenal character of a conscious experience (as Carruthers maintains), then why think that one needs to appeal to a HOT module in a creature's mind in order to explain the nature of phenomenal consciousness? It makes some sense to appeal to a HOT module if *seems red(a)* is taken to be equivalent to the *higher-order* mode of presentation *experience of red(a)*. But, if *seems red(a)* is taken to be a first-order mode of presentation, there would appear to be little motivation for accepting a HOT theory of phenomenal consciousness.

On the face of it, Carruthers' HOT account seems to face a dilemma. If *seems red(a)* is a higher-order mode of presentation, then his account runs afoul of the first-order transparency thesis; but if *seems red(a)* is a first-order mode of presentation, then there appears to be little motivation to accept a HOT account of phenomenal consciousness.

4. Conclusion

I suspect that Carruthers has replies to my three criticisms in section 3. And I hope that he does. For his project in *Phenomenal Consciousness* -- to solve *the* hard problem of the mind-body problem -- is truly a noble one, and one that I would like to see accomplished along functionalist lines. What *Phenomenal Consciousness* unarguably accomplishes, however, is a clear presentation of the strongest and most sophisticated attempt to date to solve this hard problem by means of a HOT theory. For this reason alone, *Phenomenal Consciousness*, is compulsory reading for anyone interested in the contemporary philosophical and scientific debate over the nature of phenomenal consciousness.

Notes

<1>Carruthers uses the subscript letter 'a' to indicate that these contents are analog, in contrast to the digital contents of thoughts and beliefs. Analog contents are "more fine-grained" than digital contents, according to Carruthers, enabling a subject to make subtler discriminations than he would otherwise be able to make in virtue of being in a state with digital content (p. 133-134).

<2>Carruthers recognizes that such a maneuver poses a problem for *analytic* functionalists who wish to give a functional definition of the concept of phenomenal consciousness (p. 78). But, Carruthers does not himself subscribe to analytic functionalism. Rather, he subscribes to *theory-theory* functionalism, which, unlike analytic functionalism, is not committed to "the viability of an analytic-synthetic distinction" and allows that some concepts -- recognitional concepts, in particular -- can be prized "apart from [their] surrounding beliefs and theories" in which they exist (p. 8).

<3>It should be noted that Carruthers does not deny that animals and young children possess mental states. Nor does he deny that these creatures' mental states possess *worldly* subjectivity -- that is, that there is something that *the world* is like for these creatures in virtue of their possessing mental states. Rather, what Carruthers denies is that these creatures' mental states are conscious states and hence possess *experiential* subjectivity (i.e., phenomenal consciousness).

<4> Carruthers takes Fed Dretske's FO account to exemplify the first horn of the trilemma, Robert Kirk's FO account to exemplify the second horn, and Michael Tye's FO account to exemplify the third horn.

<5> Or why conscious experiences should be any more phenomenally conscious than the experiences of shapes by visual-form agnosics. Subjects who suffer from visual-form agnosia -- as a result of damage to their temporal lobes -- seem to lack conscious visual awareness of the shape and orientation of objects, and yet they are capable of responding appropriately to the shape and orientation of objects placed in their visual fields, indicating that they are visually aware of the shape and orientation of these objects at a non-conscious level. Carruthers uses the case of the non-conscious visual awareness of shape and orientation by visual-form agnosics to illustrate the (alleged) mystery that FO theorists create by taking up option (ii). He challenges such FO theorists to explain why, on their account, "we [can not] say that the orientation of a letter-box slot *looks a certain way* to someone with severe visual-form agnosia?" (p. 171). The same (alleged) mystery for the FO theorist can be illustrated with the non-conscious experiences of blindsight subjects and absent-minded drivers, as I have done. I have decided to illustrate the (alleged) mystery for the FO theorist with these cases of non-conscious experiences only because I have described these cases earlier in the paper and have not done so for the non-conscious experiences of visual-form agnostics.

<6> In fact, Carruthers himself has come very close to using this non-higher-order sense of 'aware of' when describing the absent-minded driver case. (David Armstrong (1997, p. 723), another HO proponent, comes very close as well.) In Carruthers (1989), he writes:

While driving the car over a route I know well, my conscious attention may be wholly abstracted from my surroundings. I may be thinking deeply about a current piece of writing of mine, or phantasizing about my next summer's holiday, to the extent of being unaware *of what* I am doing [and, we can add, seeing] on the road. Yet there is a clear sense in which I must have been seeing, or I should have crashed the car. My passenger sitting next to me may correctly report that I had seen the lorry double-parked by the side of the road, since I had deftly steered the car around it. But I was not aware *of* seeing the lorry, either at the time or later in memory [my emphasis]. (p. 258)

<7> This is not to say that I cannot also sincerely say, describe, or remember *that* I am (was) seeing a red tomato. My point here is not to deny that creatures (such as adult human beings) can be higher-order aware that they have experiences. Rather, it is to show that creatures can *also* be aware *of what* their experiences represent, and that this is *not* the same as their being higher-order aware that they have experiences -- even though these two kinds of awareness may overlap in normal, adult human beings.

<8> Again, this is not to deny that blindsight subjects are unable to be higher-order aware that they are having visual experiences. Rather, the point here is merely that these subjects are *also* unable to be aware *of what* they are experiencing, and that this is not the

same as their inability (which they have) to be higher-order aware that they are having visual experiences.

<9>I defend this distinction between the higher-order and the non-higher-order reading of 'S is aware of E' in Lurz (2001) and Lurz (forthcoming). However, in that paper, I do not use the non-higher-order reading as a way of explaining why all and only conscious experiences are phenomenally conscious. Nevertheless, I see no reason why this reading could not be so employed by the FO theorist. And if it can be, I see no reason why the FO theorist cannot, then, escape from impaling himself on the point of option (ii).

<10>It should be noted that, in explaining the (alleged) mystery created by the FO theorist who takes up option (ii), Carruthers uses Robert Kirk's FO account as a case in point. As far as I can tell, Kirk does not employ the non-higher-order reading of 'S is aware of E' in his FO account; and, therefore, Carruthers may very well be correct in claiming that Kirk's FO account creates a mystery. But, the substantive question here is not whether Kirk's unique FO account creates a mystery, but whether FO accounts by their very nature create a mystery. And I do not see that Carruthers has shown that the answer to this substantive question is yes.

<11>The first-order transparency thesis is stronger than the representational transparency thesis in that it entails, but is not entailed by, the latter.

References

Armstrong, D. (1997). What is consciousness? In N. Block, O. Flanagan, and G. Guzeldere (Eds.), *The nature of consciousness* (pp. 721-728). Cambridge, MA: MIT Press.

Carruthers, P. (1989). Brute experience. *Journal of philosophy*, 86, 258-269.

Carruthers, P. (2000). *Phenomenal consciousness*. Cambridge: Cambridge University Press.

Clements, W. and Perner, J. (1994). Implicit understanding of belief. *Cognitive development*, 9, 377-394.

Dretske, F. (1995). *Naturalizing the mind*. Cambridge, MA: MIT Press.

Harman, G. (1990). The intrinsic qualities of experience. In N. Block, O. Flanagan, and G. Guzeldere (Eds.), *The nature of consciousness* (pp. 663-675). Cambridge, MA: MIT Press.

Lurz, R. (2001). Begging the question: A reply to Lycan. *Analysis*, 61.4, 313-318.

Lurz, R. (forthcoming). Neither HOT or COLD: An alternative account of consciousness. *Consciousness and Cognition*.

Povinelli, D. (1996). Chimpanzee theory of mind? In P. Carruthers and P. Smith (Eds.), *Theories of theories of mind* (pp. 293-329). Cambridge: Cambridge University Press.

Tye, M. (1995). *Ten problems of consciousness*. Cambridge, MA. MIT Press.