

ALEX BYRNE

## SOME LIKE IT HOT: CONSCIOUSNESS AND HIGHER-ORDER THOUGHTS

(Received in revised form 18 April 1995)

### I. INTRODUCTION

Consciousness is the subject of many metaphors, and one of the most hardy perennials compares consciousness to a spotlight, illuminating certain mental goings-on, while leaving others to do their work in the dark. One way of elaborating the spotlight metaphor is this: mental events are loaded on to one end of a conveyer belt by the senses, and move with the belt – perhaps changing as they go – towards a fixed circle of light, which does not completely cover the width of the belt. Some mental goings-on fail to pass through the illumination, in which case they never become conscious. But others are illuminated, and thereby enter one's consciousness. Beyond the spotlight, at the other end of the conveyer belt, lies the filing cabinet of memory, into which some of the more garish or lurid of the belt's contents fall.

This metaphor is capable of more diverse interpretations than most, and different philosophers will take issue with different interpretations. The metaphor might suggest that conscious contents are presented to an inner homunculus (what is the spotlight for, if not to show someone something?) Or it might suggest that the conscious order of mental events is the order of their arrival "in" consciousness. Or it might suggest – taking the illuminated circle to have sharp edges – that there is always a fact of the matter whether any particular mental event is, or isn't, conscious. Or it might suggest that there is a particular place in the brain where all conscious events occur.

However things stand with these four,<sup>1</sup> there is another interpretation which holds out the promise of a *reductive analysis* of consciousness: an account that gives necessary and sufficient

conditions for a mental state to be conscious in terms that do not presuppose or employ the notion of consciousness. On this interpretation, we are to think of the spot as illuminating just what mental states the subject is currently aware of being in. That suggests that what makes a mental state conscious is the fact that the subject is aware of being in that mental state.

There are two ways of developing this suggestion, depending on how “awareness” is understood. Perhaps we become aware of being in a mental state because we perceive, with some inner eye, that we are in that state. Or perhaps we just become aware by coming to believe that we are in that state. Either to be conscious is to be perceived, or else it is to be believed. Locke appears to have taken the first option, claiming that “[c]onsciousness is the perception of what passes in a Man’s own mind”.<sup>2</sup> And in our own time David Armstrong has argued that “consciousness is no more than *awareness* (perception) of inner mental states by the person whose states they are”.<sup>3</sup> (As it happens, Armstrong thinks that perception can be reduced to the acquiring of beliefs,<sup>4</sup> but of course he thinks that there is a difference between believing that the cat is on the mat and seeing that it is, and that is the only distinction we need for the purpose at hand.)

The perceptual model of consciousness held by Locke and Armstrong has peculiar difficulties of its own. Here’s one. A subject need not believe the testimony of his senses. In this familiar respect perception is unlike belief. Perception presents to the subject the world’s being a certain way, but the subject may not accept that the world is this way. Suppose then that we perceive our own mental states by some inner eye. It appears to me by my outer eye, let us say, that there is a tree before me. And, further, it appears to me by my inner eye that I am seeming to see a tree. Now I may doubt the testimony of my outer sense. It may appear to me that there is a tree before me, and yet I may believe that there is no tree before me. If we take the inner eye story seriously, I should be able to doubt the testimony of inner sense. This ought to be possible: it appears to me that I am seeming to see a tree, yet I believe that I am not seeming to see a tree. But this does not seem to be possible. It is just the familiar – and rather unhappily stated – point that there is no appearance/reality distinction for the appearances themselves. Unhappily

stated, for there is no guarantee that appearances are either incorrigible or self-intimating. That is, there is no guarantee – at least in what we’ve said so far – that when I believe I am seeming to see a tree I really am seeming to see a tree. Neither is there any guarantee that when I am seeming to see a tree then I believe I am. All that is ruled out is that it appears to me that I am seeming to see a tree, and yet I believe that I am not seeming to see a tree.

For this reason, the perceptual model of consciousness does not seem to me to be promising. So I shall concentrate here on the second option, that of construing awareness as (occurrent) belief. On this account, to adapt Armstrong, consciousness is no more than awareness (*belief*) of inner mental states by the person whose states they are. This theory is the *higher-order thought hypothesis*. The clearest and best case for the higher-order thought hypothesis has been made, in a series of fascinating papers, by David Rosenthal.<sup>5</sup> Accordingly, I shall concentrate on Rosenthal’s version of the hypothesis, and the arguments he supplies to support it.<sup>6</sup>

Before we begin, we should take note of a useful distinction drawn by Ned Block, between what he calls *phenomenal* consciousness and *access* consciousness.<sup>7</sup> A mental state that is phenomenally conscious is a state that there is something it’s like to be in. I can easily put myself in a phenomenally conscious state by stubbing my toe, opening my eyes, taking a sip of Guinness, and the like. When Thomas Nagel said that “[c]onsciousness is what makes the mind-body problem really intractable”,<sup>8</sup> he had phenomenal consciousness in mind.

An access conscious state is, roughly, one whose content is available for various cognitive operations: action, reasoning, and verbal report. Now there is nothing in general that it is like to have a conscious thought – that is, conscious thoughts need not be phenomenally conscious – so some other sense of consciousness is required. And perhaps access consciousness fits the bill.

The higher-order thought hypothesis does in fact offer a rival account of the consciousness of thoughts, but that dispute will be largely put to one side in what follows.<sup>9</sup> Our focus here will be on whether the higher-order thought hypothesis can – as Rosenthal claims – explain phenomenal consciousness in terms of certain kinds of nonconscious mental states (see TCC, 351–3).

By speaking of phenomenal consciousness, I do not mean to be taking sides on any controversial issue. Everyone believes in phenomenal consciousness, understood as I explained it. Yes, even Daniel Dennett and Paul Churchland. Of course, there are many heated areas of dispute in the vicinity. For instance, there is the question of whether phenomenal consciousness can be characterized in purely intentional terms. There is the question of whether phenomenal consciousness involves “non-conceptual content”. There is the Nagelian question of whether phenomenal consciousness is compatible with physicalism. And there are many others, but we do not have to enter into these disputes here.<sup>10</sup>

The discussion will proceed as follows. In the next section I shall explain the higher-order thought hypothesis, and examine Rosenthal’s main – and, I think, unpersuasive – argument for it. In section III I shall argue that two common objections to the hypothesis are inconclusive. In section IV I shall try to disarm a new objection, due to Fred Dretske. In section V I shall raise three other problems for the hypothesis; and finally, in section VI, I shall argue that even if the hypothesis were true, it would not satisfactorily explain phenomenal consciousness.<sup>11</sup>

## II. THE HIGHER-ORDER THOUGHT HYPOTHESIS

First, we need some Rosenthalian terminology. A *thought* is “any episodic intentional state with an assertoric mental attitude” (MD, fn. 2, 913; see also TC, 41, CSQ, 32). A thought, then, is what some philosophers would call an occurrent belief. A *higher-order thought*, or HOT, is a thought about some mental state.<sup>12</sup> To be *conscious of* something, or *transitively conscious of* something, or *aware of* something, is to be “in a mental state whose content pertains to that thing” (TC, 27); for example, to have a thought about, or to perceive, that thing. *Sensory states* are states like visual experiences, or experiences of pain, which involve sensory qualities.<sup>13,14</sup> When conscious, sensory states are phenomenally conscious. Conscious states are always *intransitively conscious*.

With a restriction to be noted in a moment, Rosenthal’s version of the higher-order thought hypothesis is simply this. “[A] mental state is conscious . . . just in case one has a roughly contemporaneous

thought to the effect that one is in that very mental state” (TOT, 199; see also TCC, 335, TC, 36–7).<sup>15</sup>

The restriction is this. Rosenthal notes that one can have a thought that one is in a mental state because of the testimony of others, or because one has observed one’s behaviour (TC, 27). But having such a higher-order thought – one based on *conscious* inference or observation – is not sufficient to make the state in question conscious. So, Rosenthal says, “[w]e must specify that our transitive consciousness of our mental state[s] relies on neither inference nor observation . . . of which we are transitively conscious” (TC, 28). I’ll mostly leave this qualification tacit in what follows.

So, for example, my thought that it’s raining will be a conscious thought just in case it is accompanied by a higher-order thought to the effect that I think that it’s raining (FPO, 30). Again, a sensory state, such as my having a headache, will be conscious just in case it is accompanied by a higher-order thought to the effect that I have a headache.

Intransitive consciousness turns out, then, to be non-intrinsic (that is, relational),<sup>16</sup> being explained in terms of transitive consciousness (specifically, higher-order thought rather than inner perception). And Rosenthal’s theory requires that (intransitive) consciousness is an accidental property of mental states, for any mental state that is the object of a higher-order thought presumably need not have been.<sup>17</sup> So, in particular, Rosenthal is committed, as he of course recognises, to the existence of nonconscious states that have sensory qualities – nonconscious headaches, visual experiences, and so forth. Disputable, no doubt, but perhaps there is nothing wrong with this idea. There is the familiar example of the headache that one is only conscious of intermittently, and Rosenthal has expended considerable energy in trying further to sweeten the pill.<sup>18</sup> These are deep and controversial waters, but fortunately for my purposes, we need not wade into them. So, for the sake of the argument, I propose to grant this starting point.

Of course we are not usually aware of the higher-order thoughts that, on this theory, make certain states of ours conscious. But this is no objection. For the theory says that a state is conscious just in case one is aware that one is in it. It does not say that we also have to be aware that we are so aware. A higher-order thought, that is, may

well not be the object of a further higher-order thought, and if it is not, then it is not conscious. Indeed, Rosenthal thinks that we rarely have such further higher-order thoughts:

Typically one's higher-order thoughts are not themselves conscious thoughts. Indeed, our feeling that the consciousness of mental states is somehow immediate is most vivid in just those cases in which the higher-order thought is not conscious. This is because conscious higher-order thoughts normally distract us from the mental states they are about, so that those states no longer occupy centre stage in our stream of consciousness (TOT, 205).

As I said, the higher-order thought hypothesis requires that no mental state is essentially conscious. The present point shows that it also requires that *some* of a subject's mental states be nonconscious. For unless the hierarchy of higher-order thoughts goes on for ever, some higher-order thoughts are not going to be the object of other higher-order thoughts, and hence will not be conscious. And since we can be sure that human psychology, at least, is not hierarchical in this fashion, then if the higher-order thought hypothesis is correct, some mental states must be nonconscious.

The point of course applies to the perceptual model of consciousness with equal force, a fact which Locke apparently overlooked. For Locke also held that every mental state is conscious, and Leibniz was not slow to point out the difficulty:

[I]t is impossible that we should always reflect explicitly on all our thoughts, for if we did, the mind would reflect on each reflection, *ad infinitum*, without ever being able to move on to a new thought. . . . It must be that I stop reflecting on all these reflections, and that eventually some thought is allowed to occur without being thought about; otherwise I would dwell for ever on the same thing.<sup>19</sup>

The higher-order thought hypothesis does, I think, have some intuitive appeal, if only for the reason that the slogan "a conscious state is one you're aware of" sounds offhand plausible. Wisely, Rosenthal does not rest his case here. His main argument for the higher-order thought hypothesis is as follows.<sup>20</sup>

First, Rosenthal notes the distinction between reporting and expressing.<sup>21</sup> If I assert that *p*, I am not only *reporting* that *p*, but also *expressing* my *thought* that *p*. So, if I assert that it's raining, I am *reporting* an external state of affairs – that it's raining – and also *expressing* my *thought* that it's raining.

Now suppose a mental state *S* of mine is a conscious state, and suppose I have the capacity to report my mental states. Then, in

virtue of the fact that *S* is conscious, I can report that I am in *S*. That is, I can express my higher-order thought that I am in *S*. So, whenever I am in a conscious state *S*, I have the ability to report that I'm in *S*, and hence the ability to express my higher-order thought that I am in *S*. It does not immediately follow from this that I am actually having, whenever I am in *S*, the higher-order thought that I am in *S*. But, says Rosenthal, “[i]t is unclear how one could have the ability to express some particular thought without actually having that thought. The best explanation of our ability to express the higher-order thought in question is plainly that one actually has that thought” (TOT, 204).

So far, we have the result that whenever I am in a conscious state *S*, I am also having the higher-order thought that I am in *S*. What about the other direction? Rosenthal argues as follows. “When a mental state is not conscious, we cannot report on it, and thus we cannot express higher-order thoughts about it. The best explanation of our inability to express higher-order thoughts about nonconscious mental states is that when the states are not conscious no such higher-order thought exists” (TOT, 204–5).

So, a state *S*'s being conscious and having the thought that one is in *S* always go together, and this suggests, according to Rosenthal, that “a mental state's being conscious consists simply in its being accompanied by such a higher-order thought” (TOT, 205).<sup>22</sup>

Of course, this is not supposed to be apodeictic, and indeed it seems pretty easy to resist. Take the first part of the argument. I am in a conscious state *S*. So (let's grant) I have the ability to report that I'm in *S*. So, I have the ability to express the higher-order thought that I am in *S*. Rosenthal says that it is unclear how I could have the ability to express some thought without actually having that thought. But it isn't at all unclear. For it may simply be that in virtue of being in *S* I have these two abilities: the ability to acquire a higher-order thought about *S*, and the ability to express it, when I have acquired it. One might sum up these facts by saying that in virtue of being in *S* I have the ability to express the higher-order thought that I am in *S*. And plainly I can have this composite ability without having the higher-order thought whenever I am in *S*. Note that to rebut Rosenthal it is not required to show that this rival explanation is *always* correct – we simply need that it might, for all we know, be *sometimes* correct.

So the argument that whenever I am in a conscious state *S* I have the higher-order thought that I am in *S* is unpersuasive. The other direction is perhaps a little more convincing. When I am in a nonconscious mental state *S* I cannot report that I am in *S*. Does this suggest that I do not have the higher-order thought that I am in *S*? Yes: surely there is no reason to suppose, unless more details are given, that I have the higher-order thought that I am in *S*. But Rosenthal requires something much stronger: that a nonconscious state *S could not* be accompanied by a higher-order thought that one is in *S*. And the argument does not show *that*.

### III. TWO STANDARD OBJECTIONS

There are two standard objections to the higher-order thought hypothesis – both raised by Rosenthal himself. Before proceeding any further, we should give these an airing.

The first is that, barring a particularly bizarre kind of panpsychism, being transitively conscious of *x* does not in general make *x* conscious. As Rosenthal says, “[m]y being conscious of a stone does not make it conscious” (TC, 30). And Ned Block complains that thinking about the states of one’s liver does not make them conscious.<sup>23</sup> So why does thinking that one is in a mental state make *it* conscious?

To this objection, Rosenthal replies that it rests on the mistaken assumption “that a state’s being intransitively conscious is an intrinsic property of that state” (TC, 31). But I do not see that this diagnosis is correct. To be sure, if you think that intransitive consciousness is an intrinsic (that is, non-relational) property of mental states then of course you will think that the higher-order thought hypothesis is mistaken, for on that account intransitive consciousness *is* relational. Rather, that intransitive consciousness is intrinsic is a plausible *consequence* of the objection, not the basis of it.

The objection is that if what makes a state conscious is the fact that it is the object of another mental state, then there is no explanation of why only *mental* states are conscious.<sup>24</sup> For other non-mental states can be the object of mental states, and they are not conscious. This objection plainly does not assume that state consciousness is intrinsic. Rather, it tries to derive an absurdity – e.g. conscious states



of the liver – from the premise that intransitive consciousness is relational in the way the higher-order thought hypothesis would have it.

The right reply proceeds in three parts. First, the higher-order thought hypothesis does not say that a state is conscious just in case the subject is *transitively conscious* of being in it. Rather, it says that a subject's state is conscious just in case she has a *higher-order thought* that she is in it (which, by definition, implies that the lower-order state is a *mental state*).<sup>25</sup> So the hypothesis does not have the consequence that states of the liver can be conscious.

This first part of the reply naturally invites a reformulation of the objection. What justifies you – the objector will complain – in building into the analysis that only *mental* states can be conscious? And the second part of the reply is simply that it is analytic that only mental states can be conscious. Is this unsatisfactory? Well, compare the relational analysis of being a brother. *X* is a brother just in case there is some *Y*, distinct from *X*, such that *Y* bears the sibling relation to *X*, and *X* is male. But what justifies you in building into the analysis that *X* is male? Same answer: it's analytic that only males can be brothers. And the reply in this case is obviously correct.

There is the third and final dialectical move to be made. When I think that my liver is in a certain state, this does not normally produce any distinctive phenomenology. Let's grant that if there were such phenomenology, this wouldn't make my hepatic state conscious, for it's analytic that only mental states can be conscious. Still, why *is* there no phenomenology?<sup>26</sup>

But the higher-order thought theorist must already make room for the distinction between phenomenally conscious states and those states – like thoughts – that are sometimes conscious but not phenomenally so. He must have an answer to the question of why there is no phenomenology associated with (some) conscious thoughts, despite the fact that, on his view, a conscious thought, and a conscious sensory state, are conscious because they are each the object of a higher-order thought. And once he has that answer, it will presumably yield an explanation of why thinking about the states of one's liver does not give rise to phenomenology. The answer, to anticipate, must be that the higher-order thought's content concerns a *sensory* state: being about a sensory state is what makes the phenomenal

difference.<sup>27</sup> It doesn't really matter here whether that reply is satisfactory. If it isn't, then we don't need to drag in hepatic states, for the distinction between phenomenally conscious mental states and those conscious in some other sense will not have been properly explained. And if it is, then the liver objection is not a problem.

The second objection is what we might call the "dog problem".<sup>28</sup> A dog, presumably, can have conscious states – pains, visual experiences, and the like. But it is quite unclear whether a dog has the conceptual resources for higher-order thoughts. If so, then we have a case of canine consciousness without higher-order thought, and hence higher-order thought is not necessary for consciousness.

Rosenthal has two possible replies to the dog problem. First, he could argue that the conceptual resources required to have a higher-order thought about a sensory state are meagre enough to suppose that dogs possess them (cf. TC, 37–40, TCC, 350–1). Perhaps dogs do not have the abilities to think about thoughts, but it is less clear that Fido has conscious thoughts, as opposed to conscious pains.

Now it is a disputed matter whether this first reply succeeds.<sup>29</sup> The second reply, although considerably more radical, seems to me to be preferable, by the higher-order thought theorist's lights: simply deny that dogs have conscious experiences.<sup>30</sup> This need not be some implausible or ad hoc maneuver, at least if we accept the starting point of the higher-order thought hypothesis. That is, remember, that no mental state – including pains and visual experiences – is essentially conscious. If this very pain I am now feeling, presumably together with many of its characteristic behavioural effects, could have been nonconscious, then it seems that our initial conviction that dogs have conscious pains was misplaced. What is obvious is that dogs are sometimes in pain. What is not at all obvious is that dogs have the higher-order thoughts required for their pains to be conscious.<sup>31</sup>

Admittedly, some would regard the availability of the second reply to the dog problem simply as a way of bringing out the absurdity of one of the initial assumptions of the higher-order thought hypothesis, namely the claim that sensory states are at best accidentally phenomenally conscious. Or it may be insisted that, even if this initial assumption *is* right, it's still risible to think that dogs are

not phenomenally conscious. But without much more argument, the dialectic at this point appears to be at a standoff.

The two standard objections to the higher-order thought theory do not refute it, but they do show the heavy burden Rosenthal's notion of a sensory state must carry. (And as I mentioned in section II, I will not be convicting the higher-order thought hypothesis on this count.) However, recently Fred Dretske has offered a new objection, which as far as I know has not received a reply. Supplying that will be our next task.

#### IV. DRETSKE'S OBJECTION

Dretske's argument against the higher-order thought hypothesis can be put as follows. Suppose I see Fred on Monday, and see him later on Friday, spending some time talking to him in broad daylight on both days. Suppose that Fred has a moustache on Monday that he has shaved off by Friday. And suppose that I do not notice that Fred has done some shaving. Nonetheless, surely I saw the moustache on Monday, and failed to see it on Friday. My conscious visual experience of Fred on Monday was different from my conscious visual experience of Fred on Friday: the Monday experience was of a moustache, among other things; the Friday experience was not of a moustache. The point is just that I am not aware that these experiences differ.

With these preliminaries made, Dretske then argues as follows:

We have just concluded that there can be conscious differences in a person's experience of the world – and in this sense, conscious features of his experience – of which that person is not conscious [i.e. aware]. If this is true, then it cannot be a person's awareness of a mental state that makes that mental state conscious. [The experience of the moustache] is conscious, and it constitutes a conscious difference between [the Monday experience] and [the Friday experience] even though no one, including the person in whom it occurs, [is] conscious of it. It follows, therefore, that what makes a mental state conscious cannot be our consciousness of it. If we have conscious experiences, beliefs, desires, and fears, it cannot be our introspective awareness of them that makes them conscious (278–9).<sup>32</sup>

If I understand Dretske's point correctly, it can easily be parried by the higher-order thought theorist. Let it be granted that there was a conscious difference between the Monday and the Friday experiences. In this particular case, I was conscious of the moustache

on Monday and not on Friday. As Dretske says, it does not follow from this conscious difference in my experiences that I am conscious *of* the difference. But neither does it follow on the higher-order thought theory. According to it, on Monday I have a higher-order thought that I am having a visual experience of Fred with a moustache (among other things), and on Friday I have a higher-order thought like my Monday higher-order thought, but without the moustache content. Those higher-order thoughts are correct: I do have such experiences on Monday and Friday. That is what makes it the case that on Monday I have a conscious experience of moustached Fred, and on Friday a conscious experience of clean shaven Fred. But plainly I can have these different higher-order thoughts on Monday and Friday, without being aware (on Friday, say) that my Friday higher-order thought differs from my Monday one.

So some fresh objections are needed.

#### V. FURTHER OBJECTIONS

The view being considered is that what makes a mental state conscious is the roughly contemporaneous occurrence of a thought that one is in that mental state. Up to now, I have been a little imprecise about the content of the higher-order thought. This now needs to be remedied.

Suppose I have a throbbing painful headache. Could that headache be only *partially* conscious? For example, could I consciously experience the headache as painful, without consciously experiencing it as throbbing? Or might it be that it's an all or nothing matter? On this view, either I have a conscious throbbing and painful headache, or else the headache is not conscious at all.

It is hard to see how the second option could be defended. Once it is granted that consciousness is not essential to mental states, and that a complex mental state may have a number of (e.g. sensory) qualities, such that a state with *one* such quality can be *wholly* conscious, then why couldn't the complex state be conscious only with respect to *some* qualities? There is no reason why not. So the higher-order thought theorist ought to hold a mental state may be only partially illuminated by the spotlight of consciousness.

But in the headache example, what makes it the case that I consciously experience the pain, but not the throbbing? Well, I am of course *aware* of the pain, but not the throbbing. That is, my higher-order thought is that I am having a painful headache (and not that I am having a painful throbbing headache). That is why I consciously experience the pain, but not the throbbing. So the content of the higher-order thought determines just what aspects of the headache are conscious.

Similar points apply to conscious thoughts. I may be thinking that the Taj Mahal serves a very agreeable Rogan Josh. Suppose this is accompanied by a higher-order thought that I am thinking about Indian food. What am I *consciously* thinking, according to the higher-order thought hypothesis? Well, I am just *aware* that I am thinking about Indian food, so that is what I am consciously thinking. Here my thought that the Taj Mahal serves a very agreeable Rogan Josh is only partially conscious. Its topic, but not its precise content, is conscious.

So, two things are mandatory additions to the higher-order thought hypothesis. First, a mental state may be less than fully conscious. Second, what makes it the case that aspect *F* of a mental state but not aspect *G* is conscious is that one is *conscious of F* but not *G*. That is, the accompanying higher-order thought is that one is in a mental state with aspect *F*.

And Rosenthal does indeed incorporate these two additions. Suppose I am looking at a bookcase containing a number of items, including a thimble. And suppose I am seeing that there is a thimble on the bookcase. Rosenthal writes:

We're seldom if ever conscious of all the detail that's represented in our sensory states, even sensory states at the center of our visual field. And the amount of detail we're conscious of often changes. When that happens, moreover, it needn't be the sensory state that changes, but only the way we're conscious of that state. The higher-order thought hypothesis explains these things. Higher-order thoughts represent sensory states in greater or lesser detail. So a higher-order thought might represent one's sensory state as being just of a bookcase with lots of things on it. But the higher-order thought might instead represent the sensory state in greater detail, as including a thimble. In the first case one is conscious of seeing the bookcase but not the thimble; in the second case one's conscious of seeing both (MD, 915).<sup>33</sup>

One thing that can happen, then, is that one's higher-order thought can fail to represent completely the nature of the mental state it is

about. In which case, the content of one's consciousness is precisely what does get represented by the higher-order thought.

This fact leads immediately to a worry. Suppose I am having a visual experience of a bookcase filled with books. The experience is replete with detail, but my conscious experience need not be. It will have precisely the content my higher-order thought says my visual experience has. So, for instance, the content of the higher-order thought could be: that I am seeing a bookcase. Or: that I am seeing two books. Or: that I am seeing something four feet from me. And so on.

But it is not at all clear that someone could have a conscious visual experience whose content is that there is bookcase, without the bookcase being consciously represented as having any color, shape, size, or being any distance from the perceiver. Likewise, a conscious visual experience of two books, with no conscious indication of how the books are spatially related to one another, or a conscious visual experience that there is something four feet away from me, with the conscious content leaving entirely open what the other properties of the thing may be, seem equally problematic. However, this may be too hasty. At least, reflection on *actual* bizarre cases, such as seeing boundaries with no color difference; seeing something as red and green all over; blindsight; blindness denial, and so forth, should suggest that there is really no objection here. Perhaps, to borrow a refrain from Dennett, we are mistaking a failure of imagination for an insight into necessity.<sup>34</sup>

So let us set this worry aside, and move on. The conscious content of a mental state is the content specified by the higher-order thought about that state. Now there is no problem here for conscious *thoughts*. My consciously thinking that it's sunny amounts to my having a higher-order thought about my thought that it's sunny, namely that I am thinking that it's sunny. And my having that higher-order thought presents no difficulty: I can easily think that I am thinking that it's sunny. But can the content of a visual experience – for instance my visual experience as I gaze on a sunny day towards the San Gabriel mountains – be captured in a single thought?

There are three potential objections here. First, the concepts that I can deploy in thought may be inadequate to characterise fully the content presented by the visual sensuous manifold, just as one may

lack the resources to describe exhaustively the content of a painting. And certainly if my thoughts are all expressible *in English*, that would seem to be right. Call this the *inexpressibility problem*.

The inexpressibility problem, I should add, does not *obviously* depend on the thesis that perceptual experiences have “non-conceptual” content. All that is required for the problem to get started is that the content of perception – whether or not it is non-conceptual content – may outrun the representational capacity of thought. But that is surely the default assumption, pending argument to the contrary.<sup>35,36</sup>

Second – temporarily waiving the inexpressibility problem – is there any reason to suppose that the proposition describing the content of my visual experience is one that I could think? For it would surely be an immensely complex thought (imagine the length of a sentence of English that expresses it). Call this the *problem of the unthinkable thought*.

Third, if I do actually have such an unwieldy thought, by introspection I ought to be able to make it conscious. At any rate, if I can’t do that then the higher-order thought theorist owes us an explanation of why not. (Recall that on Rosenthal’s view a higher-order thought is typically nonconscious, but may be made conscious by being itself the object of a further higher-order thought.) Yet when I try to become aware of my awareness of my visual experience, I do not stumble on such a monstrous thought. Call this the *problem of introspection*.<sup>37</sup>

Rosenthal’s reply to the inexpressibility problem is this. He writes: “No higher-order thought could capture all the subtle variation of sensory quality we consciously experience. So higher-order thoughts must refer to sensory states demonstratively, perhaps as occupying this or that position in the relevant sensory field” (TOT, 210; see also TC, 39–40, CSQ, 32–3).<sup>38</sup>

It is not at all obvious to me just what Rosenthal has in mind here, and he does not elaborate this idea further. But we can see immediately that there are problems. One typically succeeds in demonstratively referring to  $x$  by perceiving  $x$  (or at least by perceiving *something* – the box that contains  $x$ , for instance). Yet on Rosenthal’s view there is no inner eye by which one can perceive one’s mental states, and a fortiori one’s visual field.

Perhaps, though, Rosenthal need not have appealed to demonstratives. His basic idea seems to be that a mental designator may *directly refer* to a sensory state, as opposed to referring to it descriptively.<sup>39</sup> And a name, rather than a demonstrative, might serve the purpose.

But solving the inexpressibility problem by allowing non-descriptive reference to mental states (by either names or demonstratives) undercuts whatever reason there was to believe the higher-order thought hypothesis in the first place. Take my conscious thought that it's sunny. Because this is what I am consciously thinking, I can report that I'm thinking that it's sunny, and thus express my higher-order thought that I'm thinking that it's sunny. That is, the higher-order thought (fully) *describes* the content of the conscious thought. This kind of observation allowed Rosenthal to motivate the higher-order thought account. Now let us name the mental state that is my thinking that it's sunny, 'Alice'. There simply is no intuitive motivation for the view that my thought that it's sunny is conscious because I have the higher-order thought that I am in Alice, or because I have the higher-order thought that I am in *that* (referring to Alice).<sup>40</sup> And what goes for thoughts goes for sensory states too: "[t]he way one's transitive consciousness of the sensation represents it . . . determines how it appears to one from a first-person point of view" (FPO, 16).

So, if the higher-order thought theorist tries to solve the inexpressibility problem by claiming that the content of perception *can* be completely captured in thought, he owes us an argument for why this is so. Alternatively, if the solution invokes *non-descriptive* reference to mental states, then what we should demand is a reason other than Rosenthal's argument discussed in section II above, to believe the higher-order thought hypothesis.

As to the problem of the unthinkable thought, and the problem of introspection, the higher-order thought theorist could say that, instead of the single monstrous thought, there are *many* higher-order thoughts, which jointly yield the content of my visual experience, and which jointly make that entire experience conscious. These higher-order thoughts are small enough to be individually thinkable. And also the problem about introspection would be solved. When I am introspectively conscious, I *do* stumble on higher-order thoughts of a manageable size.

It appears that Rosenthal would accept this reply. He writes:



[W]e may need fewer [higher-order thoughts] than might at first appear. The content of higher-order thoughts may typically be reasonably specific for mental states that are near the focus of our attentions. But it is unlikely that this is so for our more peripheral states. For example, the degree of detail we are conscious of in our visual sensations decreases surprisingly rapidly as sensations get farther from the center of our visual field. It is natural to suppose that the content of one's higher-order thoughts becomes correspondingly less specific, and that a progressively smaller number of higher-order thoughts will refer to successively larger portions of the visual field (TC, 43).

But I doubt whether this proposed solution is satisfactory. It does not seem to account for the unity of my conscious experience. For when I gaze on a sunny day towards the San Gabriels, I consciously experience the relations between the blue sky, the mountains, and various buildings, for instance. I am not having a conscious experience as of mountains, and a conscious experience as of buildings, from which I may perhaps infer that the buildings are below the mountains: I just see that they are. I don't just see the *pieces* of the jigsaw, I see the jigsaw.

I am inclined to think, then, that the higher-order thought hypothesis is mistaken. But in any case, even if the higher-order thought hypothesis is true, it won't satisfactorily explain phenomenal consciousness. Making good this claim is our final task.

#### VI. DOES THE HIGHER-ORDER THOUGHT HYPOTHESIS EXPLAIN PHENOMENAL CONSCIOUSNESS?

So far, we have considered cases where the higher-order thought only partially captures the content of the mental state it is about. But here the higher-order thought is accurate, as far as it goes. What about cases where the higher-order thought gets matters wrong? Evidently, given the distinctness of the higher-order thought and the mental state it is about, such cases are possible. For instance, I may be seeing that there's a cat on the mat, and my higher-order thought may be that I am seeing that there's a dog on the mat. What would happen then? What would I consciously experience?

Well, given that I am able to report my mental states, my conscious experience is supposed to coincide with what I can report my experience to be. (At least, the higher-order thought theorist requires this plausible assumption, and we saw that Rosenthal makes explicit use of it.) Now when I report (say) that I am seeing a dog, I am

expressing my higher-order thought that I am seeing a dog. Even if I am in fact having a visual experience as of a cat, what I can report is determined by the content of my higher-order thoughts. Whether these higher-order thoughts are inaccurate is neither here nor there, as far as my conscious experience is concerned. Now what I can report in the cat on the mat example is that there is a dog on the mat, so that is what I will consciously experience. It will seem to me as if I'm seeing a dog on the mat, not a cat on the mat.

And this is Rosenthal's view, as is brought out particularly clearly in his discussion of two examples drawn from Dennett's *Consciousness Explained*. The first is this. Dennett invites us to imagine entering a room papered with identical photographs of Marilyn Monroe. According to Dennett, "you would 'instantly' see that this is the case" (354). However, you would only have time before concluding this to foveate one or two Marilyns. "We know", Dennett continues, "that parafoveal vision *could not* distinguish Marilyn from various Marilyn-shaped blobs, but nevertheless, what you see is *not* wallpaper of Marilyn-in-the-middle surrounded by various indistinct Marilyn-shaped blobs".<sup>41</sup>

Rosenthal agrees with Dennett about the content of one's visual experience when looking at the Marilyn wallpaper. "The higher-order thought hypothesis", he says, "suggests how this might happen. The relevant sensory state represents foveal shapes as Marilyns and peripheral shapes as largely indistinct shapes. One's higher-order thought, then, cleans things up [by] being a thought that one is in a sensory state in which foveal and peripheral shapes are all Marilyns" (MD, 915). So here the higher-order thought does not represent the state it is about completely accurately. It is false that your visual experience represents peripheral shapes as Marilyns. But that is what your higher-order thought claims, and so that is what you consciously experience: Marilyns all the way down.<sup>42</sup>

Dennett's second example is the phenomenon of "filling in". We do not notice a gap in our visual field corresponding to the blind spot: the area, because it's where the optic nerve leaves the retina, that lacks both rods and cones. When, for instance, a white circle on an expanse of uniform magenta falls within the blind spot, you will perceive an unbroken expanse of magenta. The blind spot has been "filled in" with magenta. And there are more dramatic effects,

such as certain sorts of pattern completion, which needn't concern us here.

According to Rosenthal, "the higher-order thought hypothesis fits well with Dennett's view that so-called 'filling in' takes place not by the brain's manufacturing the requisite sensory states, but by its forming the requisite judgements" (MD, 915).

On Rosenthal's account of filling in, the corresponding higher-order thought is again false. Your sensory state says nothing about the color of the region in your blind spot, but your higher-order thought has it that you are in a sensory state depicting an unbroken expanse of magenta, or whatever. And that is what you consciously experience, of course. Higher-order thoughts are trumps, as far as the content of consciousness is concerned.<sup>43</sup>

In the examples of the Marilyn wallpaper and of filling in, the higher-order thought in question is mistaken, but not wholly so. In the wallpaper example your sensory state represents foveal shapes as Marilyns, and the corresponding higher-order thought is right about that. And in the case of filling in, your higher-order thought accurately represents what your sensory state itself represents in the area outside the blind spot. But there is no reason, of course, why the higher-order thought should not be more seriously in error, a fact which Rosenthal explicitly recognises:

[Higher-order] thoughts can presumably occur even when the mental states that the higher-order thoughts purport to be about do not exist. But such occurrences would not constitute an objection to this account. It is reasonable to suppose that such false higher-order thoughts would be both rare and pathological. Nor would they be undetectable if they did occur. We can determine the presence of nonconscious mental states by way of their causal connections with behaviour and stimuli, both conscious and not. Similarly, we can detect the absence of mental states by virtue of the causal connections they would have with such other events (TCC, 338–9).

Suppose I have the higher-order thought that I am in a certain sensory state, and suppose I'm not in this state. Having got this far, there is only one answer to the question of what I will consciously experience: it will seem to me, phenomenologically, that I am in this sensory state.<sup>44</sup> (Consider Rosenthal's treatment of the blind spot. Imagine the blind spot enlarging to fill your entire visual field, with your higher-order thought doing progressively more filling in.)

So, for example, I could have a higher-order thought that I am in pain, but without actually being in pain. It would appear to me,

phenomenologically, that I am in pain. But the absence of pain would be detectable, for presumably I would not exhibit any pain-behaviour. Again, I could have a higher-order thought that I am having a visual experience as of a tree, but without having a visual experience of any kind. But the absence of the visual experience would be detectable, because I would not behave as if I am seeing a tree – I might walk straight into it, for instance.

In both cases I am in a state that there is something it's like to be in. That is, I am in a phenomenally conscious state. So it was a mistake to say that the higher-order thought hypothesis explains consciousness – in particular, *phenomenal* consciousness – in terms of transitive consciousness, and thus in terms of relations to other mental states. For higher-order thoughts whose content is that one is in a sensory state are not phenomenally conscious because they are the object of other mental states. It is not their relations to *other mental states* that make them phenomenally conscious. Rather, they are phenomenally conscious because they arise without any inference or observation of which the thinker is transitively conscious.<sup>45,46</sup> (Recall Rosenthal's qualification to this effect, noted above in section II.) At least, this must be the case if the higher-order thought hypothesis is correct.

Now the strategy behind the higher-order thought hypothesis, as Rosenthal characterises it, is this.<sup>47</sup> We start with an account of mental states that does not presuppose that they are conscious. We say that to be a mental state is to have (underived) intentionality or sensory properties (or both), where the notion of a nonconscious sensory property is argued to be in good standing. We then say that to be a conscious mental state is to be the object of another mental state, a higher-order thought. And if this is correct, then we have shown how to construct all conscious states from entirely nonconscious building blocks. The present problem is that if the higher-order thought hypothesis is true, higher-order thoughts that one is in a sensory state, and which occur in the right way, must be alone sufficient for phenomenal consciousness. And the question is why this should be thought to represent any kind of advance. Has *any* of the initial puzzlement surrounding phenomenal consciousness been dispelled?

If there is the appearance of progress here, it seems to me to be illusory. Perhaps the best way of showing this is to set up a dilemma for the higher-order thought theorist.

As I just mentioned, Rosenthal's official line is that having a higher-order thought that one is in a mental state is *not*, strictly speaking, sufficient for that state to be conscious. Visual scientists may tell me that I am having a visual experience, and I may believe them – that is, I may have a higher-order thought that I am having a visual experience. But this would not make the visual experience conscious. So Rosenthal adds in the requirement that the higher-order thought arises without the benefit of inference or observation of which the thinker is transitively conscious.

But surely it is completely mysterious why a state's having (or lacking) a certain *aetiology* should be the extra ingredient that turns it into a state that there is something it's like to be in. And in any case, once we are allowed to appeal to aetiology, why not do it at the level of sensory states, leaving higher-order thoughts by the wayside? It is the way that a sensory state is brought about, let us propose, that makes it phenomenally conscious. *That*, I take it, does not help to explain phenomenal consciousness, but it does just as well as the higher-order thought hypothesis.

That is the first horn of the dilemma. To get to the other horn we reject Rosenthal's concession that having a higher-order thought is never alone sufficient for phenomenal consciousness. When I am told by a visual scientist that I am having a visual experience as of a tree, and thereby come to have the higher-order thought that I am having a visual experience as of a tree, the content of my thought does not involve the right mode of presentation that suffices for phenomenal consciousness. When I am having a conscious visual experience as of a tree this is because I believe the proposition that I am having a visual experience as of a tree under a different mode of presentation – the kind that does suffice for phenomenal consciousness.

But again, if this move is allowable at the level of higher-order thoughts, it ought to be allowable at the level of the sensory states themselves. Let us say that I am having a conscious visual experience as of a tree just in case I apprehend the proposition that there is a tree before me under a certain visual mode of presentation, one that

suffices for phenomenal consciousness. Unhelpful, no doubt, but no *less* helpful than an account in terms of higher-order thoughts.

On either horn, then, the higher-order thought hypothesis cannot provide a more satisfactory explanation of phenomenal consciousness than an account that makes no mention of higher-order thoughts. And that is to say that it cannot provide a satisfactory explanation of phenomenal consciousness at all.<sup>48</sup>

So I judge the higher-order thought hypothesis to be a heroic failure. That is particularly unfortunate for me, since it is one of the few reductive accounts of phenomenal consciousness that I can understand.<sup>49</sup>

#### NOTES

<sup>1</sup> Aficionados will recognise them as some of Daniel Dennett's targets in *Consciousness Explained* (Little, Brown, 1991).

<sup>2</sup> *An Essay concerning Human Understanding*, II, i, 19.

<sup>3</sup> *A Materialist Theory of the Mind* (Routledge and Kegan Paul, 1968; reprinted by Routledge, 1993), p. 94. See also Armstrong's "What is Consciousness?", reprinted in his *The Nature of Mind and Other Essays* (Cornell University Press, 1981), and W. G. Lycan's "UnCartesian materialism and Lockean introspection", *Behavioral and Brain Sciences* 15 (1992), 216–7.

<sup>4</sup> *A Materialist Theory of the Mind*, ch. 10.

<sup>5</sup> Papers by Rosenthal are abbreviated as follows: TOT ["Thinking that One Thinks", in Martin Davies and Glyn W. Humphries (eds.), *Consciousness* (Blackwell, 1993)]; TCC ["Two Concepts of Consciousness", *Philosophical Studies* 49 (1986), 329–59]; TC ["A Theory of Consciousness", Report No. 40 (1990), Center for Interdisciplinary Research (ZiF), Research Group on Mind and Brain, University of Bielefeld; reprinted in Ned Block, Owen Flanagan and Guven Guzeldere (eds.), *Consciousness* (MIT Press, forthcoming)]; MD ["Multiple Drafts and Higher-Order Thoughts", *Philosophy and Phenomenological Research* 53 (1993), 911–18]; CSQ ["The Independence of Consciousness and Sensory Quality", in Enrique Villanueva (ed.), *Consciousness* (Philosophical Issues, 1, 1991, Ridgeview Publishing Company)]; SC ["State Consciousness and Transitive Consciousness", *Consciousness and Cognition* 2 (1993), 355–63]; HOT ["Higher-order thoughts and the appendage theory of consciousness", *Philosophical Psychology* 6 (1993), 155–66]; CSV ["The Colors and Shapes of Visual Experiences", Report No. 28 (1990), Center for Interdisciplinary Research (ZiF), Research Group on Mind and Brain, University of Bielefeld]; VET ["Why Are Verbally Expressed Thoughts Conscious?", Report No. 32 (1990), Center for Interdisciplinary Research (ZiF), Research Group on Mind and Brain, University of Bielefeld]; FPO ["First-Person Operationalism and Mental Taxonomy", forthcoming in *Philosophical Topics* (Spring/Fall 1994)]. Page references for an unpublished or forthcoming article are to the manuscript.

<sup>6</sup> For a view similar to Rosenthal's, see Norton Nelkin, "The Connection between Intentionality and Consciousness", in Davies and Humphries (eds.), *Consciousness*, op. cit., and references to other papers by Nelkin therein. For a version of the higher-order thought hypothesis restricted to an account of conscious *belief* (rather than conscious states in general), see D. H. Mellor, "Conscious belief", *Proceedings of the Aristotelian Society* 78 (1977–7), 87–101, and "Consciousness and degrees of belief", reprinted in his *Matters of Metaphysics* (Cambridge University Press, 1991).

<sup>7</sup> "On a Confusion about a Function of Consciousness", forthcoming in *Behavioral and Brain Sciences* 18 (1995) (page references to the manuscript). See also Block's "Evidence against Epiphenomenalism", *Behavioral and Brain Sciences* 14 (1991), 670–2; "Begging the question against phenomenal consciousness", *Behavioral and Brain Sciences* 15 (1992), 205–6; "Review of D. C. Dennett, *Consciousness Explained*", *Journal of Philosophy* 90 (1993), 181–93; and "Consciousness", in S. D. Guttenplan (ed.), *A Companion to the Philosophy of Mind* (Blackwell, 1994).

<sup>8</sup> The opening line of "What is it like to be a bat?", *Philosophical Review* 83 (1974), 435–50.

<sup>9</sup> A rival account, because an access theory says that a thought is conscious because it has a *dispositional* property – being available for certain cognitive operations; as explained in the following section, Rosenthal's higher-order thought theory says that a thought is conscious because it has a *categorical* property – being the object of a higher-order thought. (But see TOT, 207–11, for a discussion of a dispositional variant of the higher-order thought hypothesis.)

<sup>10</sup> Whether belief in phenomenal consciousness is as universal as I say might be cast into doubt by Block's assertion that Dennett (and Marcel Kinsbourne) assume that "there is no such thing as phenomenal consciousness" ("Begging the question . . .", p. 205). But here Block is adding to the ecumenical account of phenomenal consciousness given in the main text his own – avowedly "controversial" – view that phenomenal consciousness is "distinct from any cognitive, intentional, or functional property" ("On a Confusion . . .", p. 7). It's plain that Block is not taking this thesis to be part of the *definition* of phenomenal consciousness (see esp. "Evidence against Epiphenomenalism", p. 670).

<sup>11</sup> So the higher-order thought hypothesis does not *itself* say that phenomenal consciousness can be adequately explained by higher-order thoughts (else the sentence in the text would be contradictory); rather, the explanatory claim is a claim *about* the hypothesis.

<sup>12</sup> In some of the quotations from Rosenthal given below, I have silently replaced 'HOT' by 'higher-order thought'.

<sup>13</sup> Dennett says that "a twinge of pain or a glimpse of stocking would count as thoughts for . . . Rosenthal" (*Consciousness Explained*, p. 308), which is, in the context of Dennett's discussion, misleading. He may have had in mind Rosenthal's terminological deviation at VET, 2: "'Thought', here, is a generic term covering any sort of intentional mental state, whatever its mental attitude". But on Rosenthal's standard use of 'thought', a glimpse of stocking is an example of (the perceptual kind of) transitive consciousness, not thought. As to pain, Rosenthal does not think that pain is an intentional state, and so it is not an example of thought either. However, Rosenthal does hold that pains, and sensory states generally, "typically represent qualities and properties of one's physical environ-

ment or one's own body" (SC, 356). This kind of representation is not intentional representation. Rather, pain represents simply (and roughly) in the sense that the sensory quality associated with pain reliably indicates some state of one's body. In the terminology of Christopher Peacocke, pain has *informational* content but not *representational* content (*Sense and Content* (Oxford University Press, 1983), pp. 6–7). (Here I am indebted to correspondence with Rosenthal.)

<sup>14</sup> Rosenthal thinks that sensory qualities are non-representational properties of sensory fields (e.g. the visual field) of which we are aware when having a conscious sensory experience (see esp. CSV and CSQ). However nothing to follow will hinge on whether he is right about this. Those inclined to disagree with Rosenthal may understand "sensory qualities" in a revisionary manner, as certain kinds of intentional properties of mental states: see Gilbert Harman, "The Intrinsic Quality of Experience", in J. E. Tomberlin (ed.), *Philosophical Perspectives* 4 (Ridgeview, 1990).

<sup>15</sup> Rosenthal has his own arguments against the Locke/Armstrong perceptual model of consciousness (see TC, 32–5; CSQ, 31).

<sup>16</sup> Rosenthal uses 'intrinsic' and 'non-relational' interchangeably, and I shall follow suit. As it happens, I think that some properties (for instance dispositional properties) are both intrinsic and relational, but this disagreement will not affect the discussion.

<sup>17</sup> It will be apparent in section 6 that this reasoning is incomplete (although the conclusion – that the higher-order thought hypothesis as developed by Rosenthal requires that consciousness is an accidental property of mental states – is quite correct). See n. 46 below.

<sup>18</sup> See especially CSQ, and also Nelkin, "The Connection between Intentionality and Consciousness".

<sup>19</sup> *New Essays on Human Understanding*, trans. Peter Remnant and Jonathan Bennett (Cambridge University Press, 1982), II, i, 19.

<sup>20</sup> Rosenthal has another argument at TC, 54, but it seems to be merely the observation that the higher-order thought hypothesis can make a distinction that needs making: between having a conscious visual experience (for example), on the one hand, and consciously focusing on, or introspecting, one's conscious visual experience, on the other. The higher-order thought hypothesis accounts for the difference by saying that in the first case one's accompanying higher-order thought is non-conscious, and that in the second case it is conscious, being the object of a further higher-order thought. The higher-order thought hypothesis thus meets a condition of adequacy on any theory of consciousness, but this does not amount to a positive argument for it.

<sup>21</sup> See TOT, 200–3; VET, 12–4; TC, 56–7.

<sup>22</sup> In TCC Rosenthal claimed that "conscious mental states are mental states that *cause* the occurrence of higher-order thoughts that one is in those mental states" (338, my emphasis). Later, he retracted the causal condition (TOT, fn. 16; HOT, 159, 162). As far as I can see, nothing in the discussion to follow would be affected if the causal condition were to be reinstated.

<sup>23</sup> "Consciousness", p. 212.

<sup>24</sup> A related objection is that there is no explanation of why *I* can't make mental states of *yours* conscious simply by having the right kind of higher-order thoughts about them. This objection succumbs to the same solution I am about to suggest for the objection as stated by Rosenthal and Block.



<sup>25</sup> Leaving tacit the qualification that the higher-order thought arise without the benefit of inference or observation of which the thinker is transitively conscious.

<sup>26</sup> A repetition of the “because it’s analytic” maneuver here will of course be fruitless. The question is simply: why is there nothing it’s like to think that one’s liver is in a certain state? If it’s analytic that “phenomenology” only applies to mental states, that only shows that the question was phrased infelicitously.

<sup>27</sup> In the case of thoughts that *are* phenomenally conscious, the higher-order thought theorist should say that the corresponding higher-order thought is that one is in a state with both intentional and sensory aspects.

<sup>28</sup> See Block, *ibid.*

<sup>29</sup> For some doubts, see pp. 25–6 of the editors’ introduction to Davies and Humphreys, *Consciousness*.

<sup>30</sup> Rosenthal’s view (in correspondence) is that we do not know (and perhaps never will know) whether dogs are in conscious states, nor whether they have the resources for higher-order thoughts.

<sup>31</sup> Cf. Peter Carruthers, ‘Brute Experience’, *Journal of Philosophy* 86 (1989), 258–69. For criticism of Carruthers, and a defence of the first reply to the dog problem, see Rocco J. Gennaro, ‘Brute Experience and the Higher-Order Thought Theory of Consciousness’, *Philosophical Papers* 22 (1993), 51–69.

<sup>32</sup> “Conscious Experience”, *Mind* 102 (1993), 263–83. The quotation does not concern the moustache example (which I have adapted from Dretske). But it is clear that Dretske intends his remarks also to apply, *mutatis mutandis*, to this case. One of Dretske’s main concerns in his paper, as in a number of other places, is to argue for a sharp distinction between awareness “of things” versus awareness “of facts”. Whether or not he succeeds in this task is irrelevant to the status of his objection to the higher-order thought hypothesis.

<sup>33</sup> The thimble example is taken from Dennett, *Consciousness Explained*, p. 334.

<sup>34</sup> The first example (a particular case of cerebral achromatopsia), is discussed in David Hilbert, “Is Seeing Believing?”, forthcoming in *PSA* 1 (1994). The second (admittedly somewhat debatable) is discussed in C. L. Hardin, *Color for Philosophers* (Hackett, 1988), pp. 124–5. For blindsight, see Block’s “On a Confusion . . .”, and also Dennett’s *Consciousness Explained* (which also contains references for blindness denial).

<sup>35</sup> On non-conceptual content, see Gareth Evans, *The Varieties of Reference* (Oxford University Press, 1982), pp. 122–9, pp. 151–70, p. 227; Martin Davies, “Individualism and Perceptual Content”, *Mind* 100 (1991), 461–83, p. 462, and Tim Crane, “The nonconceptual content of experience”, in T. Crane (ed.), *The Contents of Experience* (Cambridge University Press, 1992). These theorists hold that if a subject possesses a concept, then the subject must be able to entertain *thoughts* involving that concept. If that is correct, then the inexpressibility problem *does* depend on the viability of non-conceptual content. But perhaps there is space for a position that holds that perceptual content is both conceptual and not completely capturable in thought. (Of course, making the disputes in this area at all precise requires an account of what it takes to possess a concept, and finding that is not easy.)

<sup>36</sup> One argument to the contrary would be a convincing case for Armstrong’s reduction of perceptual content to belief content, mentioned in section 1 above. But I think that Armstrong’s theory should be rejected (see, e.g., Evans, *The Varieties of Reference*, pp. 123–4).

<sup>37</sup> It's clear (I hope) that neither of these three objections assumes the controversial language of thought hypothesis. However, assuming LOT *would* make the first two objections more acute. For then there is more motivation for the distinction required by the inexpressibility problem, between the representational capacity of perception, and the representational capacity of thought. And the problem of the unthinkable thought then becomes a variant of the idea that a picture is worth a thousand words.

<sup>38</sup> In the quoted passage, Rosenthal is concerned with the "subtle variation" of non-representational properties of sensory fields (see n. 14 above). But not wishing to take sides on the controversial question of whether there are such properties (or, if there are, whether we can be aware of them), I have phrased the inexpressibility problem as one about intentional content. In my preferred formulation, the inexpressibility problem is one of capturing the "subtle variation" of *the external world* that we consciously experience. In Rosenthal's formulation, the problem is one of capturing the "subtle variation" of – as we might say – *the internal world*. If Rosenthal's formulation is preferred, that does not make the problem easier.

<sup>39</sup> Here I am indebted to David Hilbert.

<sup>40</sup> This point (about lack of intuitive motivation) is unaffected by controversies about the correct semantic treatment of names and demonstratives.

<sup>41</sup> *Consciousness Explained*, p. 354.

<sup>42</sup> An alternative interpretation (suggested by an anonymous reviewer) of Rosenthal's treatment of the Marilyn example is that the higher-order thought *changes* the sensory state, making it say "all Marilyns". I do not believe that this interpretation is correct, but in any case, all that is required for the purposes of my argument is that, according to the higher-order thought hypothesis, a higher-order thought *may* misrepresent the state it is about, whether or not any higher-order thoughts actually do so.

<sup>43</sup> See also FPO for further discussion of Dennett's examples.

<sup>44</sup> At TC, 48, Rosenthal writes that "a case in which one has a higher-order thought along with the mental state it is about might well be subjectively indistinguishable from a case in which the higher-order thought occurs but not the mental state." And at HOT, 163, he writes that "perhaps the cases just described [one of which includes having a higher-order thought in the absence of its object] would be subjectively indistinguishable from cases in which one had a higher-order thought that was properly connected to the mental-state token it purports to be about. But perhaps not; perhaps the cases just described would, from a first-person viewpoint, be experienced as in some way odd or dissonant." I am unclear why Rosenthal appears hesitant in these passages.

<sup>45</sup> Because, according to the higher-order thought hypothesis, there's something it's like to have a higher-order thought that one is in a sensory state irrespective of whether the thought has an object, it is natural to describe the higher-order thoughts themselves, rather than their objects (if they have them), as phenomenally conscious. But nothing of any importance hangs on this.

<sup>46</sup> So, as I remarked earlier in n. 17, the higher-order thought hypothesis does require that (intransitive) consciousness is an accidental property of mental states, but not wholly for the reason that a mental state that is the object of a higher-order thought might not have been. If we take phenomenal consciousness as a property of higher-order thoughts that one is in a sensory state, then the reason why a phenomenally conscious state is only accidentally so is because the higher-order

thought might have arisen by observation or inference of which the thinker was transitively conscious (in which case it would not have been phenomenally conscious).

<sup>47</sup> See esp. TC, sections III, IV, and TCC, section I.

<sup>48</sup> Suppose that some motivated way were found of denying that an “empty” higher-order thought – one with no actually existing object – could be sufficient for phenomenal consciousness. The present problem would not have gone away. All we need to set up the dilemma is – to put it loosely – that the higher-order thought may contribute some phenomenal consciousness of its own. Cases of mismatch between the higher-order thought and the sensory state it is about will suffice. Return to Rosenthal’s treatment of the blind spot. Your phenomenally conscious state is as of an unbroken expanse of magenta. This is because you have the higher-order thought that you are in a sensory state depicting an unbroken expanse of magenta. But in fact this is only partly right. Your sensory state depicts a “gappy” expanse of magenta. So there is an aspect of your phenomenally conscious state – its unbroken character – which the higher-order thought hypothesis explains simply by the fact that your higher-order thought arises in the right way. And this is just as mysterious as examples of phenomenal consciousness accompanying empty higher-order thoughts.

<sup>49</sup> Many thanks to Fiona Cowie, David Hilbert, Jim Woodward, and an audience at Cornell University for helpful discussion. I am particularly grateful to Ned Block, David Rosenthal, and an anonymous reviewer for *Philosophical Studies* for insightful comments and suggestions.

*Department of Linguistics and Philosophy*  
*Massachusetts Institute of Technology*  
*Cambridge, MA 02139*  
*USA*