

Consciousness, Higher-Order Thoughts, and What It's Like

Richard Brown

The Graduate Center, CUNY

Email: onemorebrown@yahoo.com

Web: <http://onemorebrown.wordpress.com>

Abstract

It has been 20 years since David Rosenthal first introduced the higher-order-thought-theory-of-consciousness. Though there is no consensus as to the merit of the theory, it has passed one major hurdle. It is not obviously wrong and so is on the short list of theories of consciousness which may actually turn out to be true. There is one small problem. On one hand Rosenthal maintains that thoughts lack qualitative properties. Thoughts exhibit intentionality, not phenomenality. On the other, he argues that having a higher-order-thought is sufficient for there to be something-that-it-is-like-for an organism to have any given sensory state. I argue that if this second claim is true then it *must* be the case that a higher-order-thought about, say, a belief will ensure that there is something that it is like for the organism to have the belief. In fact it turns out that all higher-order theories of consciousness entail that thoughts have qualitative character. I first give the argument that Rosenthal is committed and then show why anyone who accepts the higher-order strategy is also committed. Finally I suggest a model of the qualitative component of thought that fits nicely with the account of sensory qualitative properties that Rosenthal gives.

It has been 20 years since David Rosenthal first introduced the higher-order thought theory of consciousness and though there is no consensus in the field about the merit of the theory, it has passed one major hurdle. It is not obviously a non-starter. It has earned a spot on the short list of theories of consciousness which may actually turn out to be true. Even more surprising is that a few of the other theories on that list differ merely in this or that detail from Rosenthal's. In his new book *Consciousness and Mind* (Rosenthal 2005, unless otherwise noted all references will be to this book) he has collected several of his earlier essays, updated in the light of the preceding 20 years of argument, as well as added some new essays that develop various parts of the theory. What we end up with is a mature statement of the theory that is sophisticated, intuitive

and much more ambitious than one might think. His aim is to give a complete account of our mental life. The picture that emerges is at once powerful and seductive, able to explain a large amount of intuitive and scientific data in a simple and elegant way and there is much in the theory that I agree with. However, there seems to me to be an inconsistency between Rosenthal's general conception of the mind and the specifics of his higher-order account. It is my aim to develop this inconsistency into a problem and then offer an answer to this as yet unnoticed objection.

In brief, the inconsistency is as follows. On the one hand, he maintains, as do most of those who work in cognitive science, that thoughts lack qualitative properties. Thoughts exhibit intentionality, not phenomenality. There is nothing that it is like for an organism to have a thought. On the other hand, he argues that having a (suitable) higher-order thought is necessary and sufficient for there to be something that it is like for an organism to have any given sensory state. It is solely the higher-order thought about a sensory state, say a pain, which is responsible for there being something that it is like for the organism to have the pain. I argue that if this second claim really is supposed to be true then it *must* be the case that a higher-order thought about a first-order, say belief, will ensure that there is something that it is like for the organism to have the belief. In short, the higher-order thought theory commits Rosenthal to the claim that there is a qualitative aspect to thought.

The problem is that if you deny this, and argue, as Rosenthal does, that higher-order thoughts result in there being something that it is like for the organism only when they are directed at a sensory state then it looks like there really is something special about the sensory qualitative properties. It is not simply 'being represented by a higher-

order thought' that is responsible for there being something that it is like for the organism to have the state. But if there is more to qualitative consciousness than higher-order thoughts then the theory fails at its ambition of giving a satisfying explanation of consciousness and therefore becomes decidedly less interesting. Even worse, according to Rosenthal, is that we would then be forced to accept a picture of the mind that renders qualitative consciousness mysterious and scientifically unassailable.

Interestingly, this is not simply an argument that any one who accepts Rosenthal's version of higher-order theory is committed to qualitative properties for thoughts; this must be true of anyone who accepts what we might call the higher-order strategy. This is because all versions of higher-order theory are committed to the claim that it is the higher-order state (whatever it is) that results in there being something that it is like for the organism to have whatever first-order state they happen to be in. If this is right then all higher-order theories of consciousness entail that thoughts have qualitative character. I am going to first give the argument that Rosenthal is committed and then show why anyone who accepts the higher-order strategy is also committed. Finally I will suggest a model of the qualitative component of thought that fits nicely with the account of sensory qualitative properties that Rosenthal gives.

So, then, what is the higher-order strategy? Contemporary theories of consciousness can be roughly divided into two categories. There are those who accept the *transitivity principle*, and those that deny it. The transitivity principle is a hypothesis about the nature of conscious states that is inspired by our common sense understanding of those states. When we think about the difference between a conscious mental state and an unconscious mental state it is natural to say that an unconscious mental state is a

mental state that we are not conscious of being in. For instance it is natural to say that an unconscious belief is simply a belief that we are not conscious of having. This leads us to formulate the transitivity principle as follows. *A conscious state is a state that I am conscious of myself as being in (in some suitable way).* There are, of course, different ways that one might be conscious of oneself as being in these various states and this gives rise to the different versions of higher-order theory but they all accept the transitivity principle. Now, this is supposed to be more than just an intuitive platitude about conscious states; it is supposed to be a theory about what conscious states are. So why should we take the transitivity principle seriously as a *theory* of consciousness?

Rosenthal argues that the transitivity principle's chief virtue as a theory of consciousness lies in its ability to explain what a conscious state is (p 30). The strategy is simple. We explain transitive consciousness (consciousness of) and then we explain state consciousness in terms of the first non-problematic theory. Since the first theory is truly taken to be philosophically non-problematic, that means that qualitative consciousness would also be philosophically non-problematic. Of course this means that a lot depends on just what this non-problematic account of transitive consciousness turns out to be but if there is one, it would yield a huge payoff!¹ For the other side, that is people who think that a state's consciousness is something that is intrinsic to it, that is the end of the rode for them. Some mental states (perhaps all) are conscious and that is that. If we take this road, we are in effect admitting that consciousness is just a big mystery and throwing up our hands. What kind of an explanation is that? It is to give up the hope for a scientific explanation of this fundamental aspect of reality. So if it were to turn out that the transitivity principle could not give us an explanation of qualitative consciousness then

our reason for taking it seriously as a theory would take a severe blow. With that as background, let's look at the argument.

The transitivity principle says that a state's being a conscious state is not an intrinsic property of that state. That means that any given state can, theoretically, occur unconsciously. Thus it is a prediction of all higher-order theories that qualitative states can occur unconsciously.ⁱⁱ This, of course, brings up the question about what it means for a qualitative state to be unconscious. Rosenthal argues that if a sensory state occurs unconsciously then there is nothing that it is like for the organism to have that state. This makes some intuitive sense. If you are completely unaware of a pain state, then what sense does it make to say that you feel that pain? If you are completely unconscious of the red sensory state what sense does it make to say that it looks red to you? We understandably miss this the first time around because when we think of pains and sensations of red, we think of the *conscious* ones. So, according to Rosenthal, an unconscious pain is one that is not painful for the creature that has it. It becomes painful for the organism when the organism becomes conscious of the pain. Now, all pain states will have qualitative properties but when those qualitative properties are unconscious there will be nothing that it is like for the organism which has them. This means that it is the higher-order thought, and only the higher-order thought, that makes it the case that there is something that it is like for the organism to have the state, the first-order state plays no role in determining what it is like for the creature to have the state.ⁱⁱⁱ

But if it is simply being accompanied by a higher-order thought that accounts for there being something that it is like for the creature to have a conscious experience of pain or of red, then it should do the same thing for a belief. If this is denied, the natural

question is ‘why not? What is the difference between thoughts and qualitative states that results in there being something that it is like for a creature to have one when it is conscious but doesn’t result in there being something that it is like to have the other when it’s conscious?’ The answer to this question cannot be something like ‘well, qualitative states have qualitative properties, so when we are aware of them it is naturally like something for us to have them, but thoughts do not have qualitative properties so when they are conscious there is naturally nothing that it is like for us to have them,’ because that is to admit that there is something special about qualitative properties after all.

We were told that it was solely having the higher-order thought represent the state that accounted for there being something that it is like for the organism to be in that state but now we find out that it takes more than simply having a higher-order thought for there to be something that it is like to have mental states. We have to have a thought about *qualitative states*. What is it that is special about these kinds of states? It is their content: they have qualitative properties. So there is something about qualitative properties that makes there be something that it is like for us to have sensory states. The higher-order thought just makes us conscious of that. It is starting to look like the transitivity principle is simply just an intuitive platitude about conscious states, not a philosophical theory.

This actually sounds just like what the people who deny the transitivity principle say. Thus we lose the explanatory power of the theory by denying that thoughts are qualitative states and so lose our main reason for taking the transitivity principle seriously as a theory about what conscious states are. So if a higher-order thought makes a pain conscious and that means that there is thereby something that it is like for the creature to

have the pain then when a higher-order thought makes a belief conscious it must also thereby make there be something that it is like for the creature to have the belief. That is the argument that Rosenthal is committed to thoughts having qualitative properties but why is anyone who accepts the transitivity principle committed?

Some people find it hard to believe, to put it mildly, that a state could literally be a pain state and yet not be painful for the organism in which the state occurred. First of all how would we even know that it was a pain state in the first place if it was not painful for the organism to have it? One may then think that if they deny the claim that an unconscious pain is one that there will be nothing that it is like for the creature that has it then they can have a version of higher-order theory that can allow that thoughts lack a qualitative component. But the claim that unconscious sensory states are not like anything to have for the organism in which the state occurs is something that the transitivity principle commits you to; it is not an optional addendum.

For, if there is something that it is like for the organism to have an unconscious pain then the transitivity principle fails as an explanation of what qualitative consciousness consists in and so, again, it loses its principle motivation. The higher-order strategy is one of explaining what qualitative consciousness is in terms of some prior non-problematic theory about transitive consciousness. In order for this to work it has to be the case that whatever (suitable) means of being conscious of ourselves as being in some sensory state is solely responsible for there being something that it is like for us have that state; otherwise we have not given an explanation of the phenomenon in question.

The situation then is that the transitivity principle predicts that sensory states can occur unconsciously and if the transitivity principle is to succeed as a theoretical explanation of what qualitative consciousness consists in then unconscious sensory states must not be like anything for the organism that has them which means that it is the higher-order state, and solely the higher-order state, whatever it happens to be, that is responsible for there being something that it is like for the creature to have the first-order state in question. This in turn generates the dilemma. Either conscious thoughts are like something for the creature that thinks them or the higher-order theory cannot explain what a conscious mental state is. Anyone who accepts the transitivity principle has to face this dilemma.

So that is the argument that anyone that accepts the higher-order strategy is committed to the thesis that there is something that it is like for an organism to have a conscious thought. If it turns out that there is no qualitative aspect to thoughts, as a lot of people are inclined to believe, then it looks like all higher-order theories are dealt a fatal blow. But there are some independent reasons to think that there is a qualitative component to thoughts. Some are theoretical, like, for instance, as an explanation of how we have access to our thoughts (see for instance, Goldman 1993). In the sensory case, we appeal to qualitative properties as an explanation of how we know what conscious experiences we are having, isn't it reasonable to think that qualitative properties would also be a natural candidate for explaining how it is that we have mental access to our thoughts? Others reason are intuitive, like, for instance, reflection on introspection. There is a growing literature on hearing ambiguous utterances in one way verses another, or hearing a sentence 'comprehendingly' as opposed to as gibberish (see Pitt 2004 for a nice

introduction to this literature). Why then do so many cognitive scientists introspect and come away convinced otherwise?

The most common form of argument that I hear goes like this. I am thinking that $2+2=4$, now I am thinking that $2+3=5$ and there is no difference between them. I then think that Brooklyn is in New York and still I see no difference. Since there is no difference between these thoughts in qualitative character there is no qualitative aspect to the thoughts. The assumption here is that each distinct thought must have a distinct qualitative feel to it, and that if there were a qualitative component to thoughts we would be able to notice that difference via introspection. But there is another kind of explanation we can give for why it is that these thoughts do not seem to differ from one another when we introspect. There would be, then, no intuitive reason to think that thoughts don't have a qualitative component.

If instead of thoughts we talk about propositional attitudes this becomes easier to see. Each propositional attitude plausibly consists of two parts; namely a mental attitude directed at some represented proposition. It is natural to think, at least for propositional attitudes like hope, fear, anger, and so on, that the mental attitude itself has a qualitative component. Goldman (Ibid) offers us a nice intuition pump. Imagine a Mary-like thought experiment with a super scientist we'll call Gary. Instead of never seeing color Gary has never had certain a certain propositional attitude, say disappointment. Now imagine that he suddenly is disappointed that he will not be let out of his room. Won't he have learned something new? Namely won't he now know what it is like for him to be disappointed? It seems to me that this suggests that there is a qualitative aspect to these kinds of mental attitudes. I think we can extend this common sense idea to beliefs and desires.

To begin with beliefs, it is natural to say ‘I feel very strongly that evolution is true’, or ‘I don’t know how I feel about Intelligent Design’ which suggests that it is part of our folk psychology that the mental attitude of belief has a qualitative feel. We can also extend Goldman’s thought experiment. Imagine another scientist, Jerry, who has never had a *conscious* belief, though he has had plenty of other conscious experiences and plenty of unconscious beliefs. Now imagine that suddenly one of his beliefs becomes conscious, he becomes conscious of himself as believing something. Won’t there be a qualitative difference before and after? Won’t Gary learn something new? If this is true then having a belief, like having a desire, or being angry, consists in taking some qualitative attitude towards some represented proposition, but what qualitative property?

I have suggested that we call it the quality of conviction (Brown forthcoming). To have a belief is to feel convinced (to some degree or other) of the truth of a certain represented proposition. All beliefs are similar in this respect; all desires differ from all beliefs in exactly this respect: They feel different. Within each of these kinds of attitudes we see intensity continuums. You can be firmly convinced, just barely convinced, not quite convinced, etc. So to use an analogy belief is red, desire like green. Just like red contains all of the various shades of red, so belief contains all of the intensities of conviction and just as all shades of red are more similar to each other than they are to shades of green, so too all the various intensities of belief are more similar to each other than they are to all the various intensities of desire.

We can now see why the belief that $2+2=4$ and the belief that $2+3=5$ seem not to differ in qualitative character when introspected. Since they are both mathematical beliefs that we hold to be fairly certain, they will both have the same (or very similar) qualitative

character. It is normal that we would not be able to distinguish them from each other. It is even normal that we should miss this qualitative character when we first introspect since we have had these beliefs since we were very young and they are had very 'coolly' but there are times when we feel the conviction of our beliefs very strongly, and the idea of a belief that one no way felt convinced of would be a very strange thing.

This model of the propositional attitudes actually fits very nicely with Rosenthal's homomorphism theory of qualitative properties. Since all higher-order theories are committed to the existence of unconscious sensory states that means that they need a way to identify the sensory qualitative properties independently of our access to those properties (i.e. independent of their being conscious). This is where homomorphism theory comes in. Rosenthal begins by noting that we characterize our sensory qualities in terms of their resemblances and differences within families of properties. These families of properties are in turn specified by reference to the perceptible properties of things in the world (pp 139-140). For example we can characterize red as more similar to pink than to brown and so on. What we get from doing this systematically is a 'quality space' which is homomorphic to the quality space of the perceptible properties. Our being aware of the qualitative properties of sensory states explains how it is that we have mental access to the perceptible properties in the world. An unconscious pain state, then, will be one that resembles and differs other pain states in ways that are homomorphic to a family of perceptible properties, and via which we gain mental access to those properties. Though there may be other ways to independently specify the qualitative properties all higher-order theories need some way to do it and homomorphism theory looks promising.

To have a conscious belief is to be conscious of oneself as having a certain cognitive quality with respect to some represented content. And as we just saw, these cognitive qualities (that is the mental attitudes themselves) will stand in various patterns of resemblances and differences from each other in just the same way that the sensory qualities do. What are we to say about the actual homomorphism to perceptible properties? Is there any set of properties that the mental attitudes are homomorphic to in the way that the sensory qualitative properties are homomorphic to perceptible properties? That is, is there a set of properties that have similarities and differences amongst themselves which resemble and differ in a way that preserves the similarities and differences between the mental attitudes? This is important since we need a way to specify the attitudes apart from their qualitative component. Though I don't have time to go into detail, there is a sense in which can say yes to this; we can hypothesize that the homomorphic properties are the illocutionary forces of utterances. So the differences between beliefs that p and desires that p are homomorphic to the differences between the illocutionary force of the utterance of some linguistic item in the process of expressing the belief or desire.^{iv} Again, I do not have time to flesh this out, but this line of thought seems to be a promising way of extending homomorphism theory to cover propositional attitudes and so this account of the propositional attitudes should be very attractive to anyone who accepts a higher-order theory of consciousness.^v

References

Brown, R. (forthcoming). "The Mark of the Mental." *Southwest Philosophy Review*.

Goldman, A. I. (1993). "The Psychology of Folk Psychology". Readings in the Philosophy of Cognitive Science. A. I. Goldman, The M.I.T Press.

Pitt, D. (2004). "The Phenomenology of Cognition or What it is Like to Think That P." *Philosophy and Phenomenological Research*.

Rosenthal, D. (2005). Consciousness and Mind, Clarendon Press: Oxford.

Notes

ⁱ There are basically two kinds of answers to this depending on whether you are more influenced by a Quineian or Kripkean picture of how thought works and that the best answer is the Kripkean one. On the Kripkean model we explain how it is that we have thoughts in virtue of the causal connections between mental symbols and the world. On the Quineian picture we stick to the level of descriptive content. So this amounts to the question of whether we need to require that the first-order states cause their higher-order states. Rosenthal argues no, I argue yes.

ⁱⁱ Though this is unusual it is part of our folk psychology that these kinds of state do occur unconsciously. For instance, it is natural to think that you might be unaware of a pain even though your body posture betrays discomfort, as when seated at a desk focused on writing philosophy. There is also theoretical support for this claim. We know from priming studies, for instance, that unconscious sensory states have causal relations. In fact they seem to have the very same causal connections (for the most part) that the conscious sensory states do.

ⁱⁱⁱ Rosenthal is quite explicit about this, He argues that it often turns out that the higher-order thought represents ourselves as being in a state that is more or less fine-grained than the actual qualitative state. So, for instance, we may be conscious of a red sensory state as of a certain shade of red, or just as generally red. Exactly what it is like for the creature that has the state will depend on the exact way that the state is represented by the higher-order thought (p185).

^{iv} This even may even turn out to be an explanation of why it is that having language allows us to have more fine-grained thoughts, if we could defend the claim that being conscious of our thoughts in respect of their qualitative attitude towards some represented content gives us mental access to the properties of the language that we would use to express that thought. In the sensory case we become aware of the sensory qualities, which are the properties that mental states have in virtue of which they resemble and differ each other, and which resemblances and differences are homomorphic to the resemblances and differences that hold between the family of perceptible worldly properties. Our being conscious of these properties explains how it is that we have mental access to colors. So too in thought we become conscious of the cognitive qualities and this gives us mental access to the various kinds of speech acts we can perform. If this were the case then the cognitive qualities would be exactly like the sensory qualities and our theory of one could be used to explain the other

^v Let me briefly head off an objection. One might think that I have somehow defeated my purpose by offering the account of propositional attitudes that I have. For in the first part of the paper I argued that the higher-order theorist is committed to the claim that there is a qualitative component to thoughts in order to avoid having to say that there was something special about the qualitative properties of sensory states. But now I have argued that we should construe the mental attitudes as themselves consisting in qualitative properties. So haven't I just said that there really is something special about qualitative properties? Haven't I in effect agreed with the argument and then said 'let's attribute qualitative properties to thoughts to avoid the problem'. But this doesn't avoid the problem, the objection continues, it merely expands it because you have, in the end, not given an explanation for why the qualitative properties are special. In one sense this may be right but it does not pose any special problems. There *is* something special about qualitative properties and that is that all and only mental states have them. What makes a state a *mental* state is it's having qualitative properties, though these properties may or may not be conscious, so to say that only qualitative properties can be conscious is simply to say that only *mental* states can be conscious, which is exactly the right thing to say. This is in essence to agree with and explain the answer that Rosenthal has always given to the objection from the rock. Why is it that a higher-order thought about a mental state makes it conscious, but a higher-order thought about some state of a rock does not make that state conscious? The answer is simple. Rocks aren't mental, so they, and their states, are not the kinds of things that could be conscious. There is therefore no problem about why a higher-order thought makes a mental state conscious but not the rock's states. So too since mental states are states that have qualitative properties (which can occur unconsciously and which we have to independently specify via homomorphism theory or something that does its job) there is nothing mysterious going on here. Mental states are the kinds of things that can be conscious or unconscious and that means that they have properties that we can be conscious of ourselves as having and that results in there being something that it is like for us to have those states. So while qualitative states are special, there being special consists simply in the fact that they are mental.