ON CARRUTHERS' ARGUMENT FOR HIGHER-ORDER REPRESENTATIONALISM

Morgan Wallhagen

What makes a particular mental state conscious? According to one recently influential view, it is a creature's representing it in a particular way. Versions of this view are collectively known as Higher-Order Representationalist (HOR) theories. Despite their prominence in recent discussions of consciousness, I believe all HOR theories are mistaken, and in this paper, I hope to show why. Towards this end, I examine and reply to a recent argument, due to Peter Carruthers, in favor of HOR theories. In the process, I also advance a positive view about the contents of conscious experiences, according to which they are (typically) wholly first-order contents. On balance, then, the evidence favors a first-order theory of consciousness.

HOR Theories

HOR theories constitute one of the two main versions of representationalism about consciousness, the view that we can explain everything of philosophical interest about consciousness in terms of representations and their properties (plus facts about functional role, perhaps). According to HOR theories, a state is conscious if and only if (iff) it is a representation that is itself represented in the appropriate way. HOR theories contrast with first-order representationalist (FOR) theories, according to which a state is conscious iff it is a representation of the appropriate sort—no "higher-order" representation is necessary.

I believe that all HOR theories face serious difficulties and should be rejected. One version, the higher-order thought (HOT) theory, implies that, most likely, only humans over the age of 3 or so are conscious (Dretske 110-1).[1] I find this unacceptable. I think we can be virtually certain that many other mammals (dogs, cats, gorillas, dolphins, human infants) have conscious states (visual experiences, pains, feelings of hunger, cold, and so forth).[2] And Güven Güzeldere has convincingly argued that the other version of HOR theory—the higher-order perception (HOP) theory—either collapses into one of its competitors (HOT or FOR) or else there is no reason to believe it (792-800).[3]

I would be happy leaving it at that were it not for a recent argument from Peter Carruthers that appears to show, previous points notwithstanding, that some version of the HOR theory is true, if representationalism is true of a certain aspect of mentality (to be noted shortly). Since I take representationalism about the mind seriously, but do not want to accept a HOR theory of consciousness, I need to reply to Carruthers' argument.

Carruthers' Argument for HOR Theories

Here is Carruthers' argument:

Premise 1: There is a conceptual distinction between *worldly subjectivity* and *experiential subjectivity*. Worldly subjectivity is a matter of how *the world* seems to a subject, how the world is subjectively presented to a subject, or what the world is like for the subject (127-8). Experiential subjectivity, by contrast, is a matter of how the creature's *experiences of the world* seem to the subject, what those experiences are like, or the fact that the creature's experiences have a subjective aspect (127-8).

Premise 2: Experiential subjectivity is the distinctive, indeed distinguishing, feature of *conscious* (as opposed to non-conscious) mental states (128).

Premise 3: According to representationalism, experiential subjectivity—one's experiences seeming a certain way—is explained by a creature's representing its experiences. This follows from the fact that, according to representationalism, something's seeming a certain way is explained by the fact that it is represented a certain way (128).

Therefore, if representationalism is true of "seemings," then a HOR theory of consciousness is true. If we are to be representationalists about consciousness, we should be HOR theorists (128-9).[4]

Carruthers' argument appears persuasive. Premise 1—that there is a conceptual distinction between worldly and experiential subjectivity—is as certain as the distinction between the world a creature experiences and its experiences of that world. Premise 3 simply spells out an implication of representationalism about mentality. And there appears to be a strong case for premise 2, for it may seem to follow from "Nagel's slogan": a state is conscious iff there is something it is like for its subject to undergo it (519). If there is something it is like to undergo conscious states, does it not follow that those states seem some way to their subjects; that is, that there is something *those experiences* are like; that is, that those states have experiential subjectivity? Plausible as this may seem, I believe Carruthers' argument falters on premise 2. At least some conscious experiences lack experiential subjectivity. To see this, we need to examine Carruthers' conception of experiential subjectivity more closely.

Experiential Subjectivity Examined

In the argument above, experiential subjectivity was characterized as the fact that our experiences seem a certain way. According to Carruthers, this is just an initial characterization that needs further elaboration. Carruthers' considered view is that the difference between worldly and experiential subjectivity comes down to a difference in the *contents* of representations (183-4). A state has worldly subjectivity (it represents the world in a certain way) just in case it has (wholly) worldly content—if it (merely) attributes a property to some worldly object (the creature's body included). Examples include: representing a chair as *blue*, a ball as *red and round*, a tree as *at a great distance*, and so on. States with experiential subjectivity, by contrast, have contents that involve the property of something's being an experience. The contents of these representations thereby become, not (or not just) *world*-involving, but *experience*-involving (call such content "experiential content").

In general, a state will have experiential content iff it has content of the form: *seems x*, or *experience of x*, where "x" (typically) designates some worldly property (Carruthers 241). According to Carruthers, this experiential content "piggy-backs" on a representation's worldly content (241). In particular, a state with worldly content, *x*, acquires experiential content when its content gets added to, or modified,[5] changing from *x* to *seems x*, or *experience of x* (184, 241). For example, a worldly representation of redness acquires experiential content when it gets transformed from a representation with the content *red* to a representation with the content *seems red*, or *experience of red*. The addition of

experiential content, according to Carruthers, endows the representation with a dimension of "seeming" and hence experiential subjectivity (184). And this is supposed to explain why there is something it is like for its subject to undergo it.

Do Experiences Have Experiential Content? The Argument from Accuracy Conditions

I deny that experiential content is the distinguishing feature of conscious experiences. I contend that at least some conscious experiences have wholly worldly content. To make my case, I appeal to a very general, very basic fact about things with content (i.e., representations), namely, that they have *accuracy conditions*—conditions under which they accurately represent their objects.[6] Photographs, maps, and paintings all have conditions under which they are accurate, namely, when their objects have the properties they portray them as having. They are inaccurate otherwise. Notice, now, that some condition (something's being F) is among the accuracy conditions of a representation (of x) iff F is among the contents of the representation. For, suppose, first, that F is among the contents of some representation. Then the representation represents something, x, as F. But then x must be F for the representation to be accurate (at least with respect to x's being F). So x's being F is an accuracy condition of the representation in question. Suppose now that F is not among the contents of the representation. Then the representation does not represent its object, x, as F. But then, x's being F would not make the representation (to that extent) accurate, and thus x's being F would not be among the representation's accuracy conditions. Therefore, F is among the accuracy conditions of a representation iff F is among the contents of the representation. The correspondence between accuracy conditions and contents established, we should be able to appeal to the accuracy conditions on conscious perceptual experiences to determine whether or not they have experiential content.

Let us consider a simple experience (the complexity of the experience does not matter). Suppose you are on the beach, gazing at the clear blue sky. The sky looks blue to you. Consider your experience of the sky as it is for just a moment—say, a second. For this experience to be accurate (veridical), the sky must in fact be as you experience it, namely, blue.[7] If the sky is not blue (if it is green, or gray), then the experience is inaccurate: things are not as they seem in experience. By the correspondence between accuracy conditions and content, your visual experience of the sky has *blue* among its contents—as is obviously true. Furthermore, for your experience to be accurate, the sky must be extended through space and distant. And again, *being distant* and *being spatially extended* are among the contents of the visual experience in question. None of this is controversial.

The experience we are considering is a conscious experience, yet the contents we have thus far are purely worldly, for the accuracy conditions are purely worldly; they have to do with how the world must be for the experience to be accurate. According to Carruthers, we ought to be able to find some experiential content here, too. In particular, one's visual experience should have the content, *experience of blue* or, *seems blue*. If it does, then, for one's experience to be fully accurate not only must the *sky* be a certain way, one's *experience* must be a certain way.

I find no evidence of such content. It seems to me that the accuracy conditions on the experience of the sky are exclusively worldly. For if the sky is in fact blue, spatially extended, and distant, then this is *sufficient* for the accuracy of one's visual experience. Nothing else need be true, either about the world or oneself, for the experience to be fully accurate. To see this, simply note that it is possible for all of the accuracy conditions on one's visual experience of the sky to be satisfied—that is, for the sky (and everything else the visual experience presents) to be exactly as portrayed in that visual experience—*even if there are no experiences.*[8] For it is possible for the sky to be blue, expansive, and distant, in the way experience presents it, even where there are no experiences to present it as such.[9] Hence, one's experiences can be fully accurate experiences of a world in which there are no experiences (paradoxical though that may sound).[10] But this *shows* that the accuracy conditions, hence contents, of my experience do not include facts about, or features of, *my experience* itself. If the content of my experience did pertain to my experience in this way, then the accuracy conditions on that experience would guarantee the existence of the experience itself, and indeed guarantee that it had certain properties—just as the accuracy conditions of the experience *do* require that the sky exist, and that it be blue. Since this is not the case—since the accuracy conditions of my experience can be fully satisfied in a world where there are no experiences—the experience does not have experiential content. Conscious experiences, therefore, need not have experiential content, and hence Carruthers' claim that experiential subjectivity is the defining feature of conscious states is mistaken.

Non-conscious Mentality and Worldly Subjectivity: A HOR-theorist Reply

There is an interesting way Carruthers could reply to my argument. As Carruthers notes, there appears to be good evidence for the existence of non-conscious mental phenomena of many kinds, including non-conscious sensory states (147-68).[11] Such states, according to the representationalist, represent (portions of) the organism's body and ambient space. These representations, therefore, possess worldly content—hence worldly subjectivity—even though they are non-conscious. But then, the argument goes, the possession of worldly content is not sufficient for a state's being phenomenally conscious. Perhaps, then, the representationalist must accept the existence of experiential content to account for the distinction between conscious and non-conscious mentality.

Of course, the final step does not follow. It does not follow from the fact (if it is a fact) that worldly content is not sufficient for a mental state to be conscious that experiential content *is* what makes them conscious. It does not even follow that experiential subjectivity exists. Still, the argument may seem to retain some force. For one thing, some might think it obvious that the most reasonable thing to infer is that the distinguishing feature of conscious states is experiential content.[12] After all, many philosophers have the intuition that conscious states must differ in their contents from non-conscious states. And minimally, one might think, the argument shows that possession of worldly content cannot be sufficient for a state's being conscious.

But in fact, the argument does not even establish this much. For, there may be different *types* of worldly content. If so, then even if possession of *merely worldly* content does not render a mental state conscious, it may yet be true that possession of a *type* of worldly content does. This would still avoid commitment to the existence of experiential content.

Two Types of Worldly Content: Phenomenal and Non-phenomenal

I believe the first-order theorist can offer a compelling version of the view that there are two kinds of worldly content. Simply put, the view is that worldly content comes in phenomenal and non-phenomenal forms. Non-conscious states have non-phenomenal worldly content; conscious experiences have phenomenal worldly content.[13]

The argument for the distinction between phenomenal and non-phenomenal worldly content stems from reflection upon the accuracy conditions of (otherwise similar) conscious and non-conscious states. Let us consider, then, a conscious experience of a red vase and a non-conscious experience of the same red vase. It will be helpful, first, to note the distinction some philosophers draw between phenomenal color and physical color. Physical color is a property of physical surfaces, a property, perhaps, in virtue of which those objects cause us to have experiences of color. But physical color, unlike phenomenal color, need not be (and presumably is not) *similar to* color as experienced. Indeed, we can understand the claim that modern science portrays the world as "colorless" as the claim that physical surfaces do not possess properties similar to colors as we experience them. Phenomenal colors, by contrast, are similar to colors as we experience them.

In terms of the distinction between physical and phenomenal colors, I claim that conscious vision represents objects as phenomenally colored.[14] My conscious experience of the red vase, for instance, represents the vase as being phenomenally red. If the vase does not have that property—phenomenal redness[15]—then there is a sense in which my experience is inaccurate, a sense in which my conscious experience presents the vase as having a property it does not have. If a state's having phenomenal content amounts to its having, among its accuracy conditions, an object's possessing some phenomenal property, then conscious visual experiences have phenomenal content. And note that the accuracy of this experience depends on what property *the vase* has. The visual content in question, the phenomenal content *red*, is thus worldly content.[16]

But what of non-conscious experiences? Imagine a non-conscious representation of the red vase. It may help to adopt, for the moment, the view of some HOR theorists, Carruthers included, that humans are the only animals with conscious experiences. These theorists, of course, do not deny that many animals develop visual representations of their environments. Indeed, they accept that some of these animals, certain monkeys for instance, have color vision in the sense that they develop representations of the colors of objects. Suppose such a creature is perceiving our red vase. And suppose we accept the view, suggested by modern science, that the world is not phenomenally colored, and, hence, that the vase is not phenomenally red. Is the creature's non-conscious visual representation of the vase in any way inaccurate? I see no reason to think it is. But, by contrast, our conscious visual experiences *are* inaccurate if that picture of the world is true, and this is precisely because of their phenomenal content. So there is a strong case for the view that non-conscious states have *non-phenomenal* worldly content.

Since the distinction between phenomenal and non-phenomenal worldly content is well motivated, the existence of non-conscious states with worldly content is not an

insurmountable objection to first-order theories of consciousness. The first-order theorist can account for the distinction between conscious and non-conscious states (in terms of phenomenal vs. non-phenomenal content), maintain that (some) conscious states have purely worldly content—and hence deny that experiential content is their defining characteristic—while nevertheless allowing that non-conscious representations have worldly subjectivity. Furthermore, we can now see that the phenomenal/non-phenomenal distinction *is not the same as the worldly/experiential subjectivity distinction.* I believe Carruthers' basic error is to suppose that these distinctions are the same. They are not. Conscious states can be distinguished from non-conscious states by the fact that they are phenomenal even if they lack experiential subjectivity.

Note that it will not help to protest that all Carruthers *means* by "experiential subjectivity"—hence such locutions as "seems x," "experience of x," and so forth—is, in my terms, phenomenal character. For example, phenomenal-red is a potentially worldly property, a property that objects could have. Experience-of-red is *not*. Again, the accuracy conditions on a representation of *phenomenal-red* make no reference to experiences, but the accuracy conditions on a representation of *experience of red* (obviously) do. So, they are not the same notion.[17]

I emphasize this point because it undermines one of Carruthers' most important claims on behalf of HOR theories, namely, that they can explain phenomenal consciousness where first-order theories cannot. For, once we see that phenomenal character and experiential subjectivity are distinct, we can see that, for all Carruthers says, HOR theories do not explain consciousness either. So, Carruthers has neither established a HOR theory of consciousness nor shown that such theories have any advantage over first-order theories.

Conclusion

I have argued that some conscious experiences have wholly worldly phenomenal contents. If so, experiential subjectivity is not the defining characteristic of conscious states, and Carruthers' argument for HOR theories fails. More positively, my account of the contents of conscious experiences suggests a first-order theory of consciousness, since those contents are first-order. Combine this with the facts that HOR theories have no clear advantage over first-order theories (as just noted), and have positive disadvantages (noted toward the beginning of the paper), and the balance of evidence favors a first-order theory of consciousness.[18]

NOTES

1. HOT theorists are well aware of this objection, but reply to it in different ways. See Rosenthal 741-2; Carruthers 193-208.

2. Many people have this intuition, of course, but it is not *merely* an intuition, for a wide variety of considerations—behavioral, anatomical, and evolutionary—can be brought forth in support of it. See Griffin 28-100, 142-53.

3. I urge readers who are not familiar with Güzeldere's paper to read it, especially if you are inclined to believe a HOP/HOE/Inner Sense theory of consciousness.

4. So put, the conclusion of Carruthers' argument might appear to be of interest only to representationalists. This bothers me little, since many philosophers interested in consciousness are in fact inclined to take the representationalist claim about seemings/appearances seriously. Still, the argument should be of general interest. For the argument purports to show that a certain kind of representationalist theory is true if *any is*. As such, the argument, if good, tells the philosopher who is not persuaded of representationalism which representationalist theory she needs to argue against.

5. It is not entirely clear which best captures Carruthers' view about what happens when experiential content comes into being.

6. Searle often notes this, though his preferred phrase is "conditions of satisfaction" (119). Some philosophers speak of "truth conditions" in connection with content, but this is a mistake, since not all representations are true or false: some exhibit a different kind of property, veridicality, or accuracy. Maps are one example. A map can be more or less accurate, but is not thereby true or false. I think of truth as a particular kind of accuracy or veridicality.

7. Though it helps in thinking about the example, one need not assume that there is an actually blue object—or any object at all—picked out by "sky" for the example to work. Indeed, it only makes my point to note that there is no object that corresponds to how we experience the sky. For this is generally noted in the context of pointing out how our experience of the sky is misleading—i.e., inaccurate in a way.

8. What about the sky's seeming to be very distant? This content (hence accuracy condition) does require the existence of a being, a creature that the sky is very distant from, but that is all. The creature need not be experiencing the sky.

9. It is important to note that the view that qualities such as color are observer dependent does not present a problem for the present argument. For even if that view is correct, the naïve conception of colors as mind-independent properties of objects in the world (i.e., naïve realism) is at least *coherent*. As such, it is coherent to suppose that the world would be as portrayed in my experience even if I did not exist—even if no experiences existed at all.

10. What *is* impossible is to experience a world un-experienced. For when one speaks of a world as "un-experienced" one speaks of a world in which the act of experiencing has not occurred. For similar reasons what cannot see what is unseen, think what is unthought-of, imagine what is unimagined, etc. The mental acts in question make the things seen, thought-of, imagined, etc. Nevertheless it is possible to experience, think, imagine a world *as* unexperienced, unthought-of, unimagined. And this is because doing these things is a matter of undergoing mental states with particular contents, contents that do not include the fact that the acts themselves are occurring.

11. For a good recent review of this evidence, see de Gelder, de Haan, and Heywood.

12. Even then, *something* needs to be said to blunt the force of my argument from accuracy conditions, but I set that aside for now.

13. I believe something like this view is implicit in the work of many FOR theorists. It needs to be made explicit here, to be brought to the surface and defended, to reply properly to the objection I have considered.

14. I do not intend this to imply that they do not also represent them as physically colored: perhaps they do.

15. This is a property the vase *could* have, after all, whether or not we believe it actually has it (or whether any object is actually phenomenally colored). To doubt that it *could* have such a property is to doubt the very coherence of our commonsense conception of colors. Whether or not our commonsense conception of color is correct (I doubt it), I know of no considerations that would show it to be incoherent.

16. I point this out in case anyone worried that the explicit introduction of the notion of phenomenality here created problems for the idea that the contents are worldly. I think it should be evident that, in my earlier arguments about the accuracy conditions of conscious experiences, I was presupposing that the contents of those states were what I am now calling "phenomenal contents." Those arguments would be in serious trouble if we found, here, that phenomenal contents could not be worldly.

17. It is worth noting that, for similar reasons, equating "phenomenal red" with "experience of red" (etc.) would rob Carruthers of an argument for his HOT theory of consciousness. For Carruthers believes—and I agree—that to have a representation with the content *experience of x* one must possess the concept EXPERIENCE (or something like it; 215, 219). But since phenomenal red (etc.) is a potentially worldly property, there is no reason to assume that a creature must have the concept EXPERIENCE (etc.) to have a representation with the content *phenomenal x*.

18. Acknowledgements:I would like to thank Gary Hatfield, Joel Martinez, Gary Purpura, and Michelle Wallhagen for helpful discussion and comments on earlier versions of this paper. A version of this paper was presented at the 2006 meeting of the New Mexico/West Texas Philosophical Society in Las Cruces. I am grateful to those present for their comments. I owe special thanks to Jennifer Noonan for her insightful comments, I regret that I could not take her points into account here.

WORKS CITED

Carruthers, Peter. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: MIT P, 2000.

de Gelder, Beatrice., de Haan, Edward., and Heywood, Charles. *Out of Mind: Varieties of Unconscious Processes*. Oxford: Oxford UP, 2002.

Dretske, Fred. *Naturalizing the Mind*. Cambridge: MIT P, 1995.

Griffin, Donald. *Animal Minds*. Chicago: Chicago UP, 1992.

Güzeldere, Güven. "Is Consciousness the Perception of What Passes in One's Own Mind?" *The Nature of Consciousness: Philosophical Debates*. Eds. Ned Block, Owen Flanagan, and Güven Güzeldere. Cambridge: MIT P, 1997. 789-806.

Nagel, Thomas. "What Is It Like to Be a Bat?" *The Nature of Consciousness: Philosophical Debates*. Eds. Ned Block, Owen Flanagan, and Güven Güzeldere. Cambridge: MIT P, 1997. 519-27.

Rosenthal, David. "A Theory of Consciousness." *The Nature of Consciousness: Philosophical Debates*. Eds. Ned Block, Owen Flanagan, and Güven Güzeldere. Cambridge: MIT P, 1997. 729-53.

Searle, John. *Mind: A Brief Introduction*. Oxford: Oxford UP, 2004.