COMPUTATIONAL ARCHITECTURE AND THE CREATION OF CONSCIOUSNESS

SCOTT BROCKMEIER
Arizona State University

THERE IS NOTHING so intimately familiar yet so inexplicable as conscious experience. Despite their familiarity, judging between states that we would call conscious and those we would not can be difficult. Are you conscious when dreaming? How about when hypnotized? Are you conscious of the stop sign you obey despite being so lost in thought that you have no recollection of stopping for the sign? There are even distinctions between different types of consciousness, such as Block's distinction between phenomenal and access consciousness.[1] There are, however, few theories that venture to formulate any kinds of necessary and sufficient conditions for consciousness.

David Chalmers introduces several "psychophysical principles," but these are principles intended to provide guidance in the formation of a theory of consciousness rather than constitute a theory or part of a theory.[2] Daniel Dennett argues for a Multiple Drafts Model (MDM) of consciousness.[3] The MDM seems to explain many of the often perplexing phenomena associated with consciousness. This model, however, is more of a rough outline than a concrete theory. There is not even an attempt to give sufficient conditions for consciousness. The MDM seems to be on the right track with much to offer in the development of a theory of consciousness as do Chalmers's psychophysical principles. As important and intriguing as this work is, it provides no theories of sufficient conditions for consciousness. There are, however, at least two theories that do purport to provide sufficient conditions for consciousness.

In this paper, I will examine David Rosenthal's higher-order thought (HOT) theory of consciousness as a description of consciousness in a connectionist system, viz. our brains. I will also examine John Pollock's Oscar project as an example of a serial machine with purported consciousness. My reasons for choosing these particular theories are that they both claim sufficient conditions for consciousness and have strikingly similar outlines. That one describes consciousness in a connectionist system and the other in a serial system provides a framework for discussion of the differences between serial and parallel systems.

In the first section, I will briefly describe Rosenthal's and Pollock's theories. In the second section, I will give an account of why both Pollock and Rosenthal resist thinking of consciousness as an inner per-

ception. Section 3 consists of a comparison of the two theories.

In section 4, I will interpret the impact of *phenomenal* and *repre - sentational superposition*, as described by Dan Lloyd, on the two theories under discussion.[4] Lloyd uses similarities between characteristics of connectionist networks and characteristics of human conscious experience to argue for the use of connectionist systems as models to facilitate understanding of consciousness.

Buttressed by the strength of these similarities, I will argue in section 5 against the possibility of consciousness in a machine that uses serial computation.

1. Two theories of consciousness

Higher-order thought theory[5]

Rosenthal's higher-order thought theory trades on a distinction he makes between transitive and intransitive consciousness. Transitive consciousness is consciousness *of* something. Here 'to be conscious of' is a transitive verb phrase requiring a direct object. A creature capable of consciousness could be said to be transitively conscious of what the direct object refers to. Intransitive consciousness does not involve this transition from noun phrase to direct object. This use compares to verb forms such as 'to run' or 'to sleep'. It involves no direct object. Only mental states can be intransitively conscious whereas we can be transitively conscious of both the mental and the physical. On this theory, a thought (first-order mental state *m*) becomes conscious only when accompanied by a higher-order-thought (HOT) of the form 'I am in mental state *m*'. This makes *m* conscious (intransitively) but the HOT itself is not conscious. In order for the HOT 'I am in mental state *m*' to become conscious, it needs to be accompanied by a third-order mental state which has that HOT as its object. This third-order mental state would be of the form 'I am having the thought—I am in mental state *m*'. Third-order states, though not difficult to form, are relatively rare outside philosophy or psycho-therapy. Forming successively higher-order mental states becomes very difficult very quickly.

The first-order mental state *m* could be any of a variety of types of mental states: perceptual, cognitive, emotional, etc. Let us use, as an example, the perceptual state of seeing a stop sign. On this view, one could be (transitively) conscious *of* the stop sign without the perceptual mental state itself being (intransitively) conscious. It is only when *m* becomes the object of a HOT, that is to say one is conscious *of m*, that we say *m* is conscious. It seems strange at first to say that one can be conscious *of* something while that mental state (the being conscious *of* state) itself is not conscious. Consider, though, the familiar scenario in

which you are sufficiently conscious of a stop sign to stop your car yet not so conscious of the sign that you can remember making the stop at the end of your drive. You were preoccupied with thoughts of weekend activities when you stopped at the stop sign. At the time you stopped at the sign, your plans for the weekend were (intransitively) conscious, yet you had to be (transitively) conscious of the sign in order to obey it.

Rosenthal includes a provision in his theory for HOTs arrived at through conscious inference or observation. An example of such a HOT might in the following scenario: some friends point out that you are acting uncharacteristically, that you are snapping at people and that you have a sour look on your face. They say that you must be angry about something, and you realize that what they say about your behavior is true. You conclude based on this information that you are in fact angry, yet you feel no conscious anger. Rosenthal believes that this is a case in which you have a HOT of the form 'I am in state *m*' but that state *m* is not itself conscious. To guard against such cases, Rosenthal includes the provision that a HOT arrived at through conscious inference or observation need not result in the corresponding first-order state's becoming conscious.

Oscar

According to John Pollock, his Oscar project will result in the world's first artificial intellect, or artilect. Pollock also claims that Oscar will be conscious, and thus he will presumably be able to enjoy his status as the world's first artilect. Pollock describes his Oscar project's ultimate objective as "the construction and modeling of a general theory of rationality" (Pollock 113). My focus in this paper will be on what is purported to provide Oscar with consciousness, or the experience of qualitative feel.

Pollock does not think that consciousness will simply arise from the complexity required to produce intelligence. He will instead install certain sensors that he claims will endow Oscar with consciousness. There will be a category of sensors that will let Oscar know what is occurring in the world. They will allow him to sense his environment. These are 'perceptual sensors'. The second category of sensors will be sensors trained on Oscar's internal world. Call these 'introspective sensors'. There are two types of introspective sensors. The first type are called 'first-order introspective sensors' and the second type 'second-order introspective sensors'. The first are sensors that monitor the output of the perceptual sensors. They will enable Oscar to respond to his environment according to the dictates of his goals and reasoning apparatus.

The first-order introspective sensors let Oscar know what his perceptual sensors are reporting. According to Pollock this is more than a

mere analog to qualitative feel—this *is* qualitative feel. But Oscar is not quite conscious. These first-order introspective sensors provide Oscar with qualitative feel, but in order for this feel to be conscious, Oscar must not only *have* the qualitative feel but also *experience* that feel. To *experience* the qualitative feel requires second-order introspective sensors. These are sensors trained on the first-order introspective sensors. A more detailed description of these comes later. The output then of the first-order introspective sensors becomes conscious when a second-order introspective sensor "looks" at it. This sounds surprisingly similar to Rosenthal's HOT theory with the second-order introspective sensors doing the work of the higher-order thought. This point is explored in detail below.

2. Against a perceptual model

Until fairly recently (within the last fifty years or so), most characterizations of consciousness used comparisons to some type of illumination. Consciousness as a spotlight trained on a thought or sensation was a common metaphor. Thinking of consciousness in this way appeals to intuition and is especially attractive to one who has a Cartesian view of mind-brain interaction. The two theories examined here urge us to resist such conceptions of consciousness.

Rosenthal warns against thinking of the HOTs on a perceptual model such as vision or hearing. He says that if our HOTs are perceptual like vision then there must be some distinct 'mental quality' associated with the perception. When we see a cat with our eyes, the sensation of the cat and the cat itself are distinct. Let's call the sensation of the cat *C** and the cat itself *C*. Now consider the case of a first-order perception of a dog, call it *T*, and the corresponding higher-order perception of *T*, which we will call *T**. It is not clear that these two, *T* and *T**, can be considered distinct with respect to their sensory content. The sensation is the same even in introspection, when both *T* and *T** are *in*transitively conscious. Rosenthal says that you *do* have the first-order thought of a dog and, in addition, another thought that results in the first-order thought becoming conscious. However, while the thoughts themselves are to be considered distinct, the sensation of the dog, it seems, is the same in both the first and second-order thoughts. (The only difference between the two, in the case of ordinary non-introspective consciousness, is that one is conscious and the other not.) The fact that the sensation of an external object and the object itself are distinct while the purported 'perception' of a mental state, such as a thought, results in no such distinct sensory content is meant to show that, as Rosenthal says, "the analogy with perception [is] idle" (Rosenthal 32).

Just as Rosenthal resists a perceptual model of consciousness,

Pollock resists characterizing what gives Oscar consciousness as perceptual in nature. Pollock talks of introspective sensors but this is misleading in that the term 'sensors' names the components that pick up stimuli from the outside world (i.e., perceptual sensors). The processes involved, though, are fundamentally different. In distinguishing between the perceptual and introspective sensors Pollock says:

> We talk loosely about introspection as the "mind's eye", but no one thinks there is literally a sense organ in the mind surveying what is occurring there. When we talk about incorporating introspective sensors into Oscar, we should not think of them on that model. The addition of introspective sensors consists only of treating the output of his perceptual sensors in new ways, giving them additional functional roles in cognitive processing. (Pollock 5)

By this description, there really is no sensing going on when Oscar introspects; rather, Oscar uses the output of his perceptual sensors in his cognitive processing.

Later, in *How to Build a Person*, Pollock indicates that the introspective sensors mentioned above are not sufficient to provide Oscar with consciousness. Pollock uses a distinction between *having* a qualitative feel and *experiencing* a qualitative feel. These introspective sensors provide Oscar with the former but not the latter. Oscar has the feel through the operation of his first-order introspective sensors, but in order to experience the qualitative feel, second-order introspective sensors are required. To support the view that one can have a qualitative feel yet not experience it, Pollock distinguishes between sensations and the feels of sensations. He says that to equate a sensation and the feel of a sensation would be like equating a rock and the feel of a rock.[6]

The output of the first-order introspective sensor is the sensation itself and the operation of the second-order introspective sensor results in the feel of the sensation being experienced by Oscar. Pollock points out that the output of the first-order introspective sensor is always available but usually not utilized and therefore usually not conscious. It is only when the second-order introspective sensors are "trained" on the output of the first-order introspective sensors that the feel of the sensation is experienced. Only then is the feel of the sensation conscious.

As quoted above, Pollock says that we are not to think of the first-order introspective sensors as a sense organ surveying the contents of the mind. He says that there really are no new sensors added. We merely treat the output of the perceptual sensors in new ways. They play new functional roles in Oscar's cognitive processing. If we are to think of the first-order introspective sensors as the utilization of the output of the perceptual sensors in new ways, then how are we to understand the operation of the second-order introspective sensors? How do they *sense*

the *output* of the first-order introspective sensors? Pollock writes,

> To have the feel is for one's first-order introspective sensors to be operating, but to experience the feel is to attend to the feel itself, and that consists of the operation of second-order introspective sensors sensing the operation of the first-order introspective sensors. (Pollock 14)

Pollock is not explicit about how this is to be interpreted. One possible interpretation is that the sensing performed by the second-order introspective sensors is meant to be understood along the lines of the sensing of the first-order introspective sensors. That is, perhaps the output of the perceptual sensors take on further functional roles in Oscar's cognitive processing.

Under this explanation, however, it does not seem that the second-order introspective sensors are really "sensing the operation of the first-order introspective sensors." The second-order introspective sensors would only be giving different functional roles to the same data as the first-order introspective sensors, that data being the output of the perceptual sensors. If this is the case, then it seems quite possible for the second-order introspective sensors to operate independently of the first-order introspective sensors. Oscar would be able to experience the qualitative feel while not even having that feel. This cannot be right.

Another possible explanation is that the second-order introspective sensors, since they are to sense the output of the first-order introspective sensors, operate as the first-order introspective sensors do. That is, they look to the functional roles that the first-order introspective sensors give to the output of the perceptual sensors and give this output new functional roles. I can see how a "sensor" trained on another sensor can be thought of as "sensing" the other sensor not as a true sensing but as giving the output of that sensor new functional roles. But it is harder to see how an even higher-order sensor can be said to sense, in the same (serving functional roles) way, the output of a sensor whose operation is to be understood as giving a lower order sensor's output new functional roles. How do you give functional roles to functional roles?

The operation of these second-order introspective sensors is what gives Oscar his putative conscious experiences. How they accomplish this is not clear. If Rosenthal's and Pollock's theories are both correct and describe the same thing then there must be a relationship analogous to that of the HOT and the first-order thought between the second- and first-order introspective sensors. One or both of the theories, though, may be wrong. Another possibility is that they are both right yet no definitive parallels can be drawn between what produces consciousness in us and what produces consciousness in Oscar. Perhaps the two theo-

ries describe two disparate mechanisms that produce similar results.

3. ROSENTHALAND POLLOCK COMPARED

Both Rosenthal and Pollock warn against thinking of consciousness as the result of some perceptual process. Both also have a multilevel model of what is required to produce consciousness. But are they both saying the same thing, only in different ways? The operation of Pollock's perceptual sensors would correspond to the operation of our various sense modalities. Our eyes and ears produce neural outputs. The production of these is often guided by cognitive processes but is not dependent on them. Light striking the retina will result in a signal being sent to the occipital cortex even when higher functioning of the brain is absent. Production of the signal sent to the occipital cortex does not depend on its being part of any higher cognitive processes. Similarly, in Oscar the perceptual sensors produce signals whose production does not rely on higher cognitive processing.

The operation of the first-order introspective sensors in Oscar would correspond to what Rosenthal calls transitive consciousness. Pollock says that these first-order introspective sensors are little used most of the time but they are nonetheless always operating. This seems to correspond closely to our being *conscious of* many things where only some small subset is intransitively conscious. John Searle makes what I see as the same point in a discussion of the center and periphery of consciousness. There are sensations such as the "feel of my shirt against my skin" or the "tightness of my shoes" that it would be wrong to call unconscious.[7] Searle uses the word 'conscious'differently, as compared to Rosenthal, but I believe the phenomenon is the same, albeit with divergent descriptions. So far it seems as if the two theories *are* saying

Figure 1[9]

Rosenthal	**Pollock**
- cat	- cat
- neural signals sent to brain from various sensory modalities	- perceptual sensors produce output
- info from sense processed to some degree and is transitively conscious	- operation of first-order introspective sensors provides Oscar with qualitative feel
- HOT results in some first-order thought(s) being intransitively conscious	- operation of second-order introspective sensors provides Oscar with the *experience* of qualitative feel

the same thing. *Figure 1* shows my formulation of the analogous components of the two theories. We now turn to consciousness with qualitative feel, or intransitive consciousness.

The HOT theory says that a mental state is conscious iff it is accompanied by an associated HOT that is not the result of conscious inference or observation. Pollock says that Oscar *experiences* the qualitative feel produced by the first-order introspective sensors iff a second-order introspective sensor is trained on the output of that (some of those?) first-order introspective sensor(s). The HOT then would serve the same role as the operation of the second-order introspective sensors.[8] There are questions about just how the second-order introspective sensors in Oscar operate. Let us assume that there is a coherent way to explain the operation of these introspective sensors at work in Oscar. We will find that there are other problems that will prevent us from ascribing consciousness to Oscar.

4. PHENOMENAL SUPERPOSITION AND SERIAL COMPUTATION

The reason that Oscar cannot be conscious is that he is implemented on a serial computer rather than as a connectionist system. I do not believe that programs run on a conventional serial computer can be conscious. What Oscar will lack (and what I see as a necessary feature of consciousness) is what Dan Lloyd calls *phenomenal superposition*. Our conscious experience of visually perceiving a car is an experience of a certain model, a certain color or colors, a certain type (sedan, truck, or sports car), to name just a few of the aspects that make up the perception of a car. All of these features are simultaneously present yet each is distinct. There is no blending or blurring of features. Part of the perception of the car could be that it belongs to some particular person, or that it is a car that makes you feel good because you learned to drive in that same model. These sensory, cognitive, and emotional aspects are all part of the same experience. They are experienced not as a list of features but as a single integrated whole.

Some connectionist networks display a characteristic similar to phenomenal superposition. Dan Lloyd describes *representational superposition* as "the simultaneous activation of multiple dimensions of information within a single activation vector." Representational superposition is demonstrated in connectionist networks constructed by Jeffrey Elman, which he calls "Simple Recurrent Networks."[10] One such network was fed ten thousand two- or three- word sentences formed from a list of 29 words. The sentences were given one after the other with no punctuation, e.g., 'woman smash plate cat move man break car boy move girl eat bread dog move mouse mouse move book'. The entire list was presented to the network six times, one word at a time, with the network being trained, by way of a feedback process, to

predict the next word. The network never attained perfection in its ability to predict the next word, but that was not the goal.

What was of interest was the way the net, after training, represented the 29 words that were presented. An activation vector, in a connectionist network, is the number or times each node in the hidden layer is active during processing of a particular input—words, in this case. Examination of the activation vectors of these 29 words showed that they were grouped according to various characteristics. The activation vectors of verbs were different than those of the nouns. The activation vectors for inanimate nouns were distinct from those for animate nouns. As a subclassification of animate nouns, humans were distinct from animals. The net displays representational superposition in that an activation vector of, for example, the word 'cookie' represents it as a noun that is inanimate and also as food. (To say that the vector of the word represents it as food is to say that the vector is most similar to the vectors of other words on the list that also refer to food.) When we have an experience of a car, it is a single experience, rich in phenomenal characteristics. When one of these simple recurrent nets "experiences" a word, such as 'cookie', this "experience" is rich in grammatical characteristics.

I claim that phenomenal superposition is a necessary feature of consciousness and that Oscar will not have it because he is implemented on a serial machine rather than a connectionist network. What remains to be shown is that serial computation cannot produce phenomenal superposition.

5. An argument against consciousness via serial computation

Obviously, the experience of a connectionist network, as described above, is nothing like any experience had by a person. The appeal of connectionist architecture as a model for machine consciousness is, however, compelling. One can point to the successes of neural nets, such as proficiency in pattern recognition, but even stronger is the intuitive appeal of such systems as possible candidates for consciousness. When I try to imagine what consciousness (or some analog to consciousness) in a serial computer would be like I imagine something that, if it exists, must differ not in degree but in kind. I may be accused of suffering from philosopher's syndrome[11], mistaking a lack of imagination for an insight into necessity, but I hope to show that the differences are profound enough to justify my conclusions.

I do not mean to argue that it is purely the complexity or the electro-chemical nature of the brain that allows us to be conscious. What I want to demonstrate is how fundamentally different the brain is from a computer serially manipulating symbols in binary code. How could a serial computer produce the simultaneous and superposed, seamlessly

integrated rich visual and cognitive content intrinsic to the perception of a red sports car? A program running on a serial computer might have a command that says, "Seamlessly integrate all of the information from visual experience *x*," but all of the information would still be "seen" by the computer as a single file succession of discrete bits of data.

Consider the grand finale of a Fourth of July fireworks display. Suppose that there is one simultaneous burst of 35 fireworks rockets that starts and ends within a couple of seconds. We all know what this would be like for ourselves phenomenologically. What would it be like—what could it be like—to experience such a display as a serial computer, to experience the display as succession of discrete points of light, one after the other? The device would first receive the information on one point of light as to its position, intensity and duration, then information on the next and then the next and so on.[12] Not only would the information be presented on one point of light at a time, but the information on each point would be, in effect, a single file succession of ones and zeroes. A serial machine could not reproduce the simultaneity of the fireworks experience, no matter how fast the machine ran. Even if the serial computer processed all of the information of the fireworks display in *less* time than the actual time taken for the display, the "experience" would still be, not of a single event, but of a string of temporally distinct information-laden bits of data.

It follows that any simulation of a connectionist system on a serial machine will fail to be conscious. This is true regardless of the complexity of the net being modeled or the precision of the model. What is needed for consciousness is for the entire connectionist system, or at least a significant portion of it, to be active all at once. It is only at the time when such a system is active, when signals are coursing through the net, that there could be true conscious experience. Only then can there be the kind of phenomenal superposition that is required for any state we would want to call conscious.

One might argue that, although the information is presented serially, the system's subjective experience is of a simultaneous presentation. It might seem to the system that all of the bits of information that the fireworks display is composed of are experienced together at a single point in time. This line of argument, though, leads directly to dualism.

To say that serial computation can seem to happen, subjectively, all at once is to say that there is some thing beyond the computation, some entity, that "remembers" and "experiences" data that are no longer present. This problem is not encountered in connectionist computation. In a connectionist system there is no need to postulate the existence of some thing beyond the computation that holds information until the system is ready to experience some delineated block of data. In a connec-

tionist system, all of the data *is* experienced at once. All of the information *is* simultaneously present. Consciousness in a connectionist system can only occur at the time the relevant signals are working through the system. A connectionist system *at rest* will *not* be conscious.

There is a project ongoing at MIT[13], run by Rodney Brooks, that has as one of its goals the production of a conscious robot. In this project, they are using 64 separate microprocessors connected in a custom parallel design. Such a configuration might seem to bridge the gap that I have been describing between serial and connectionist processing. One might think that although there is no massive interconnectivity, as in connectionist systems, there are multiple streams of processing being performed simultaneously and therefore there could be some crude form of consciousness present.

This type of parallel processing is still just a series of serial computations with some limited interaction. If some group of serial microprocessors were massively interconnected such that the system worked as a connectionist network, it would no longer be correct to describe the operation of such a system as serial computation. I do not know if such a scenario is technically possible. If it is possible, I do not know where we would draw the line between a serial microprocessor doing serial computation as opposed to working as part of a connectionist net. Such a line may be a very fuzzy line with some degrees of consciousness emerging along a continuum between strictly serial and strictly connectionist architectures.

Since only hard-wired connectionist systems can be conscious, Oscar will never be so endowed unless he is realized in a connectionist system. But what about the HOT theory? Is the HOT theory a possible description of what it would take to produce consciousness in a machine? Talk of a hierarchy of thoughts in a connectionist machine will require some explication of what exactly a thought is. In a human brain there is a constant base level of neural activity. Presumably, a thought is some organized pattern of activity that exceeds this base level. In a connectionist system described as conscious due to a process as described by the HOT theory, there would need to be some corresponding, organized pattern of activity. Perhaps the thought would be the activity that is present as the input signal makes its way through the net. This, though, presents the problem of the net “holding” a thought. The thought might be repeatedly recycled through the net in a cyclical pattern. Perhaps this is exactly what it is for a human to “hold” a thought. Perhaps what seems to be the continuous presence of a thought is really just a recycling of the thought, a continuous refreshing of the thought. As more sophisticated connectionist systems are developed, we may find that we *can* implement Rosenthal’s HOT theory into such a system producing a conscious machine.

The neural nets that we have for study now are quite crude, and drawing parallels between these nets and actual human neural activity in any but the most general terms is premature. Given the primitive state of brain modeling by connectionist systems, it is difficult to see how a connectionist architecture could realize conscious thought through a higher-order thought mechanism. A possible alternative is that there is a type of superposition relationship, similar to phenomenal and representational superposition, in the activation vectors of a neural net that explains consciousness. Perhaps consciousness itself is superposed as just another piece of information not unlike redness or shape.

Answering questions such as these will become easier as more complex connectionist systems are developed. Existing connectionist systems are certainly too primitive to support any kind of hierarchical structure of thoughts necessary for consciousness according to the HOT theory. As more sophisticated nets are developed, we may be able to incorporate such hierarchies and thus produce consciousness. At this point, however, we can only be certain that any machine that does acquire conscious states will not be a machine that operates serially. Connectionist systems seem likely candidates for machine consciousness, although technical problems associated with producing an actual hardwired connectionist system of sufficient complexity may prove extremely difficult.

There are interesting questions about what would result from simulating serial computation on a connectionist system. If the algorithm that is Oscar were to be simulated exactly on some vastly complex connectionist system, would this result in a conscious Oscar? I think not, for reasons similar to those given against serial computation. If Oscar does turn out to be a rational agent, answering questions like this would help to illuminate differences between mentality *qua* reasoning and mentality *qua* consciousness. This intriguing topic, however, must wait for another time and place.

Arizona State University
Tempe, Arizona

NOTES

Acknowledgements: I wish to thank Bernard Kobes for his invaluable guidance. I also thank the Sun Angel Foundation for their support.

1. See Block (1995). "On a confusion about a function of consciousness," *Behavioral and Brain Sciences*, 18, pp. 227-287.

2. See Chalmers (1995). In this paper, Chalmers presents two non-basic principles (the principle of structural coherence and the principle of organizational invariance) and one basic, lawlike (albeit speculative) principle (the double aspect theory of information). These principles are described in

detail in his book *The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford University Press (1996).

3. See Dennett (1991).
4. See Lloyd (1996).
5. See David Rosenthal (1996).
6. It seems odd to talk of a qualitative feel that is not experienced or a sensation that is not felt. Using these terms in this way is presumably a technical usage rather than a common one.
7. See John Searle (1992), *Rediscovery of the Mind*, MIT Press. The relevant discussion is on pp. 137-139. The comparison here is a little sticky in that Searle does not use the word 'conscious'in the way we have been using it. Searle frames his discussion in terms of degrees of consciousness rather than making a transitive/intransitive distinction. He does not make a distinction between different types of consciousness. He believes that everything is conscious—just conscious to different degrees, even at the periphery. I see it as the same phenomenon just without the sharp boundaries. I tend to favor some degrees of consciousness but feel that, at some point in the periphery, *conscious* fades to merely *conscious of*.
8. I say here that the *HOT* serves the same role as the *operation* of the sensor. The former seems to pick out an object, the latter an event. Pollock (1989) makes the point that his theory is about mental events as opposed to mental objects. He makes no ontological claims about mental objects. Rosenthal, on the other hand, formulates his theory using thoughts and their interrelations. Any theory that postulates the formation of, and/or causal roles of, mental objects presupposes mental events. The converse, however, is not true. One can have a theory of mental events that is not committed to the existence of mental objects. Is there a fundamental disparity between the two theories based on a dependence on the existence of mental objects? I don't see that there is. I see nothing in the HOT theory that commits Rosenthal to any particular ontological status of thought, beliefs, desires, etc. The talk of thoughts can easily be seen as an easier and more natural way to talk about mental events without any ontological commitments about mental objects.

   The token constituency requirement, though it seems to commit one to the existence of mental objects, on closer inspection does not. For a description of token-constituency see Kobes (1995) "Telic Higher-order Thoughts and Moore's Paradox." You might think that if the HOT is to have, as a constituent part, the first-order thought, then it is, obviously, necessary for the HOT to have parts. It does. But an event *can* be divided into parts. A presidential election is an event, and it makes perfect sense to talk about what happens in a particular precinct as separate and distinct from the election as a whole.
9. I take some liberty here characterizing Rosenthal's view as distinguishing between neural signals sent from the sense modalities and processing that results in transitive consciousness. As far as I know, Rosenthal never makes this distinction. It seems, though, as if there *must* be some cortical process-

ing of the signal from, say, the eye for there to be consciousness of whatever that signal represents.

10. See Jeffrey Elman (1990).

11. This phrase was coined by Daniel Dennett.

12. The information could be organized differently, such as all of the positions first, then all of the intensities, then the colors, etc., but that would not make any difference to the point being made here.

13. See Daniel Dennett (1995).

REFERENCES

Chalmers, D. "Facing up to the Problem of Consciousness." *Journal of Consciousness Studies* 2.3 (1995).

Dennett, D. *Consciousness Explained*. Boston: Little, Brown & Company, 1991.

———. "Cog: Steps Toward Consciousness in Robots." *Conscious Experience.* Ed. Thomas Metzinger, 1995.

Elman, J. "Finding Structure in Time." *Cognitive Science* 14 (1990): 179-211.

Kobes, B. "Telic Higher-order Thoughts and Moore's Paradox." *Philosophical Perspectives* 9 (1995).

Lloyd, D. 1996. "Consciousness, connectionism and cognitive neuroscience: a meeting of the minds." *Philosophical Psychology* 9 (1996): 61-78.

Pollock, J. *How to Build a Person*. Cambridge: MIT Press, 1989.

Rosenthal, D. "A Theory of Consciousness." *The Nature of Consciousness: Philosophical Debates.* Eds. Ned Block, Owen Flanagan, and Guven Guzeldere. Cambridge: MIT Press, 1996.

Searle, J. R. *The Rediscovery of the Mind*. MIT Press, 1992.