REPRESENTATIONAL THEORIES OF CONSCIOUSNESS, INTROSPECTION, AND INNER-SENSE

John O'Neal

Introduction—Representing and Introspecting

Many philosophers of a naturalistic bent have recently become enthusiastic about the possibility of explaining consciousness in terms of representation. Such accounts purport to explain how a conscious mental experience (such as a feeling of pain) is experienced in presumably non-mental things like toes rather than in the head (the presumptive residing place of the mental, according to most naturalists). According to the representationalists, our experiences are essentially sensory representations of our environment as being a certain way. One can experience a pain in a phantom limb because the mind represents the limb as being in pain. However, these attempts to reduce conscious states to perceptual representations face a potential problem. In addition to carrying information about the body and world, our minds possess introspective states that carry information about the mind itself. This is problematic for accounts that reduce consciousness to perceptual representation for there is no obvious perceptual system (or inner-eye) that allows the mind to peer in on itself.

In this paper, I will argue that the most satisfactory resolution to this problem is the adoption of the classical notion of an inner-sense. However, the conception of inner-sense I endorse differs importantly from many current accounts of inner-perception. Higher-order theorists of consciousness such as Armstrong, Lycan, and Lormand, all posit an inner-sense that is constitutive of consciousness. According to their views, a mental state becomes conscious by being introspected. I will argue inner-sense need not and should not be seen as constitutive of consciousness. I will argue that introspection is best accounted for within a first-order theory of consciousness that (contra the wishes of many first-order theorists) accepts the existence of an inner-sense/perception.

Introspection—The Explananda

Contemporary accounts of introspection need not be as ambitious as those of earlier philosophers, as introspection is now thought to be crucially limited and even fallible. Due to the contributions of Freud and many others, possessing mental states we are unaware of is uncontroversial. Not only do we now think that introspection is limited in this way, many also acknowledge (see Nisbett and Ross) that introspection may often be highly unreliable; the act of introspection itself may fabricate the existence of past mental states to rationalize our previous behavior, rather than report about what is going on in the head. However, while a theory of introspection need no longer be a complete theory of the mental, introspection's limitations and fallibility alone provide some data a theory of introspection ought to be able to explain.

Another introspective phenomenon in need of explanation is the famous "coming to" effect often discussed by David Armstrong—often when driving or walking, time passes while the mind wanders. Then we suddenly "come to" the realization that we have been maneuvering for some time without paying attention. Thus it seems we are able to (1) attend

to our experience as the happening of an experience, and (2) gain immediate knowledge concerning our occurrent mental states.

Introspection is also considered a source of knowledge about our own propositional attitudes. My knowledge about what I think seems qualitatively different from the knowledge other people can have about me. They infer it, while I experience it (or seem to). Introspective access may be neither complete nor infallible, but it does not seem to be nil either.

I propose the above considerations sketch the explananda of a minimal account of introspection. An account of introspection ought to tell us (1) why introspection is limited, (2) why it is fallible, (3) how we can come to attend to our experience, (4) how we can gain immediate knowledge our occurrent conscious states, and (5) how we come to know (at least some of) our propositional attitudes.

Perception Accounts

One way to account for our awareness of our mental states proposes we become aware of mental states in the same way we become aware of tables and chairs; we perceive them. Introspective mental states are simply perceptions of other mental states. Such an account is a higher-order perception (HOP) theory of introspection. The notion of an inner-sense capable of discerning states of the mind has both classical and contemporary defenders. Indeed, John Locke once proposed all perception is mediated by inner perception. "Consciousness is the perception of what passes in a man's own mind" (bk.2, ch.1, sect.19). Similarly the contemporary philosopher David Armstrong states "consciousness . . . is a perception-like awareness of current states and activities in our own mind" (61). Note that Locke and Armstrong identify consciousness generally with inner perception. I do not believe such introspective accounts of *consciousness* actually do well in accounting for *introspection*. Nevertheless, if there is an inner-eye, it is one way to account for our awareness of our own mental states.

And an inner-perception theory of introspection can do much of the work called for above. The view can account for introspective fallibility by pointing to a general analogy with perception. As our visual system is fallible to systematic illusions like the Mueller-Lyer effect, so might the inner-eye be fallible in systematic ways. The view may also give some insight on our introspective ability to attend (to "come-to") by analogizing with external perception. For instance, when hitting a curve ball our visual system can come to dominate our consciousness as we "block out" sound, taste, touch, etc. Similarly, the "coming-to" effect could be the result of the inner-eye's operation coming to dominate our conscious mental resources. In addition to fallibility and attention, an inner perception account of introspection will grant the same immediacy to introspective knowledge that can be accorded to any perceptual knowledge. Just as we can immediately recognize a coffee cup, we can immediately recognize our occurrent mental states.

Where the account may come up short is in its ability to explain the limitations of introspection. If there is an inner-eye peering around the mind, why is so much of our mental activity (our sub-conscious mental processing) closed off to it? Additionally it is not clear how an inner-sense can account for our knowledge of our propositional attitudes, for it is not clear what it could mean to perceive a propositional attitude.

Higher-Order Thought Accounts of Introspection

Another way of accounting for introspection is to deny any symmetry with outer perception and claim introspection is essentially a kind of belief, or higher-order thought, or fact (rather than object) awareness. As Sydney Shoemaker puts the view:

> Our minds are so constituted, or our brains are so wired, that, for a wide range of mental states, one's being in a certain mental state produces, under certain conditions, *the belief* that one is in that mental state. This is what our introspective access to our own mental states consists in. (emphasis added 268).

As such accounts associate introspection with a kind of thought or belief about a mental state rather than a perception, they can be classified as higher-order thought (HOT) theories of introspection. The idea is that along with certain first-order perceptual states—*this soup is hot*—other beliefs come immediately for free—*I am having a hot soup experience.* The details concerning how these "free beliefs" are doled out differ—some talk of accompaniment (Shoemaker and Rosenthal), others of a special kind of immediate inference or "displaced perception" (Dretske, *Naturalizing*, "Awareness"), and still others talk of inference guided by a "folk" theory of mind (see Gopnik). Regardless of the details, the end product is supposed to be something that is more proposition-like than perception-like.

HOT accounts of introspection may be able to explain the limitations of introspective knowledge better, because where our folk theory is silent or where the premises needed for a chain inference are missing, introspection will be silent. Additionally they may be able to handle propositional attitudes better, for it is natural enough to expect a chain of inference to end in some proposition.

Against HOT Theories of Introspection

While perhaps handling these two facets of introspection better, I believe there are considerations that argue decisively against a HOT theory of introspection. The first concerns a problem often brought up in discussions of David Rosenthal's HOT theory of consciousness (see Carruthers, *Consciousness*). Namely, our introspective abilities appear to outstrip our conceptual store. I can not only introspect upon the fact that I am now perceiving a curve ball, but also that I am perceiving a curve ball of just that shade, speed, etc. If what individuates an introspection is some thought I come to form about by my experience, then it seems I would need a different concept for every introspectible feature of the target experience. But not only does it appear I do not have a different concept for every introspectible feature of my experience (no "Fastball-429" concept), the idea that I could seems to imply an excessive nativism in that it puts concept possession prior to our ability to become aware of a novel experience. Indeed, it would seem to require the possession of a very large store of concepts before one could introspect anything.

Perhaps the reply of those who think introspection is a kind of belief will argue that introspective beliefs are "course-grained". The idea being that they need not carry all of the information of the perceptual states they are about. I would reply that regardless of the grain, if introspection really is belief formation it would appear we must possess the concept of orgasm (for example) before we could know we had one—which seems counter-intuitive at a minimum.

Another reply from those who see introspection as a kind of belief formation might be that the only conceptual content they require is an indexical element—i.e., they are essentially of the form "I am having *this* experience." But this seems to turn all introspection into the same introspection, and not a very useful one. If introspection is really just coming to a belief that always has the same conceptual content then it does not appear to be any real source of knowledge. Additionally this reply may actually require the existence of an inner-eye to pick out the referent of the indexical. What is the "this" in such a belief supposed to pick out? What could it point to other than the experience itself? But this is precisely what HOT theories of introspection say we do not have access to. As Dretske would put it—we are aware with our experiences, not of them.

In Defense of Inner-Sense

If we cannot live without it, can we live with the inner-eye? I believe so, but I also believe that it is important to distinguish an inner-sense theory of introspection from the inner-sense theories of consciousness that inner-sense defenders often want to push. The claim that the operation of an inner-eye can explain consciousness generally is independent of the claim that an inner-eye explains our ability to become aware of some of our mental states. It is this second claim that I want to defend.

To understand why it is important to pull these notions apart, consider two of the most influential objections to positing an inner-eye. The first concerns the so-called "transparency" of introspection (see Tye 30-1). Simply put, introspection seems to lack any unique phenomenology while our other sensory modalities carry phenomenal properties arguably unique to that sense. While sight putatively conveys the uniquely visual qualities of the sighted thing, introspection carries no obvious introspective qualities of the introspected thing (i.e., an experience). If asked to introspect upon a particular conscious experience, such as a pain in the toe, the only things available to attend to are one's toe and its pain, rather than any sort of uniquely experiential properties. Thus, inner-sense objectors claim that if introspection is a sense, it is a very odd one, for there appear to be no sensory qualities for it to track. The second concern associated with the notion of an inner-sense is its potential commitment to an infinite regress involving Cartesian theater. Is an inner-sense theorist committed to existence of a set of inner inner-eyes to perceive our introspection?

My claim is that neither of these objections carries weight for an inner-sense theorist who also maintains a first-order theory of consciousness—i.e., the idea that a state is conscious simply in virtue of its place in our cognitive architecture. A first-order theorist of consciousness like Tye holds that a state is conscious when it is the output of a perceptual module and poised (or available) to our higher-order decision processes. Contrary to this view, higher-order theorists of consciousness generally contend that a conscious state is a perception of a perception. Typically, first-order theorists (see Dretske) have found talk of inner-eyes anathema as these are often pushed as the producers of higher-order conscious mental states. But this strikes me as a mistake by both sides. If the inner-eye is simply one of many perceptual modules of the brain, then it seems perfectly coherent to say that its output will be a first-order state whose output can be "poised" in the same way as the output of our other perceptual modules. Think of it as a perception of the brain rather than a perception of a perception.

What does the inner-eye perceive and represent? I suggest it may be the activity of the other perceptual modules of the brain. An introspection can be thought of as representation of the brain as perceiving. That we do possess representations of internal perceptual states (like hunger) is uncontroversial given any representational theory of the mind. My suggestion is that as part of the brain may be devoted to representing states of the stomach, so may another part of the brain be devoted to representing the perceptual processing areas of the brain.

Restricting the monitoring duties of the inner-eye to our other perceptual systems shows a way to deal with the two theoretical shortcomings discussed in my initial account of a HOP theory of introspection. Recall that higher order perception accounts of introspection were potentially unable to account for the limitations of introspection. According to my suggestion, we should expect our introspective abilities to be limited to just those mental states that are involved with perceptual processing areas of the brain. This would explain why we cannot introspect the operation of the kinds of heuristics (primacy, recency, right-hand-rule, etc.) we use unknowingly in decision making; they do not engage the perceptual processing areas of the brain.

Such an account may also be able to explain why some propositional attitudes are introspectible and others are not. According to the above suggestion, a belief or other mental process ought to be available for introspection to the degree that its formation and possession involves some processing by the perceptual areas of the brain. That some thinking does reuse the resources of our perceptual systems has empirical support. For example, the visual system seems to be implicated in imagination (see Kosslyn). It has also been hypothesized that our auditory system may be implicated in the formation of certain judgments through the phenomenon of inner speech (see Carruthers, "Thinking"). According to the theory on offer, the degree to which a particular thought does borrow from the perceptual processing resources of the brain will also be the degree to which it is introspectible.

Restricting the domain of the inner-eye's monitoring duties, along with rejecting the attempt to explain consciousness generally, also may resolve the two traditional objections to inner-sense canvassed above. Consider the objection that the notion of an inner-eye implies a homunculus involving Cartesian theater subject to an infinite regress. Now if one agrees with Locke and Armstrong that the operation of an inner-sense can explain consciousness generally, then one may well have a problem accounting for introspection without positing further sets of inner-eyes. For if the light of the inner-eye is what brings about a conscious state, then our conscious states will not be introspectible unless there is a further eye to cast light on the workings of the first. This strikes me as just another way of making the point that introspective accounts of consciousness are not accounts of introspection.

However, if one combines a first-order theory of consciousness with an inner-sense theory of introspection, then no extra set of eyes is needed to account for the fact that our introspections are conscious. Being a conscious representation is simply being the output of a perceptual system that is available to our higher-order reasoning. Introspections are conscious for the same reason any experience is conscious—because they are the output of a perceptual system that has the function to represent a part of our environment (i.e., the other perceptual processing areas of the brain) as being a certain way (as perceiving).

My suggestion may also be able to help with the so-called transparency objection to the postulation of an inner-eye. Recall that according to this objection experience as such cannot be the object of a perception for it possesses no unique phenomenology; there are no properties for it to track. As Peter Carruthers puts the complaint, if introspection is a kind of perception then "surely we would expect there to be a *distinct* (non-worldly) set of

properties of phenomenally conscious experiences on which introspection could concentrate" (*Consciouness*, emphasis in original).

But, again, I think this objection is only damaging to those who hold the inner-eye responsible for phenomenal properties in general. Admittedly, if the attention of our inner-sense is what confers painfulness upon an experience of pain, then it is justified to wonder why the workings of the inner-eye confer no further properties upon examining its own output. But according to my suggestion inner-sense does not confer qualia; it does not make painful experiences painful. Inner-sense simply represents ("senses") that our other sensory modules are sensing. Rather than having a *unique* phenomenology or set of phenomenal properties, we should only expect introspection to have *some* phenomenology as every other sensory module does. But I think we can all attest to the fact that introspecting upon a pain or an itch or a tickle (for example) *does* in fact have a phenomenology. It makes the pain, itch, tickle more intense. And indeed, it is very hard to see how this could be possible if introspection were just a belief that one was in pain rather than something like a "first-order" perception.

Conclusion

Whether or not there is inner-sense is obviously an empirical question. Nevertheless, I hope I have illustrated research on theories that involve an inner-eye are better motivated than some philosophers have thought, and that even if scientists find an inner-eye one day, the debate between higher-order and first-order theories of consciousness will remain.

WORKS CITED

Armstrong, D. "What Is Consciousness?." *The Nature of Mind*. Ithaca, NY: Cornell UP, 1980. 55–67.

Carruthers, P. "Conscious Thinking: Language or Elimination?" *Mind and Language*, 1998. 13.

———. *Phenomenal Consciousness: a Naturalistic Theory*. Cambridge: Cambridge UP, 2000.

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

———. "The Mind's Awareness of Itself." *Philosophical Studies* 95 (1999).

Gopnik, A. "The Illusion of First-person Knowledge of Intentionality." *Behavioral and Brain Sciences* 16 (1993).

Lycan, W. "Consciousness as Internal Monitoring." Eds. Block, N, O.J. Flanagan, and G. Güzeldere. *The Nature of Consciousness*. Cambridge, MA: Bradford Books/MIT P, 1997.

Locke, J. *An Essay Concerning Human Understanding*. Indianapolis, IN. Hackett, 1995.

Lormand, Eric. "Shoemaker and Inner-sense." *Philosophical Topics*, 2001 <http://www-personal.umich.edu/~lormand/phil/cons/shoemaker.htm>.

Nisbett, R. and T. Wilson, T. "Telling More Than We Can Know." *Psychological Review* 84 (1997): 231-295.

Rosenthal, David "Consciousness and Metacognition." *Metarepresentations*, (2000):265-295.

Seager, William. "Introspection and the Elementary Acts of Mind." Forthcoming in Dialogue, <http://www.scar.utoronto.ca/~seager/mywork.htm>.

Shoemaker, S. "Self-Knowledge and 'Inner-sense.'" *Philosophy and Phenomenological Research* 54 (1994):249–314.

Tye, M. *Ten Problems of Consciousness*. Cambridge, MA: MIT P, 1995.