A "Quasi-Sartrean" Theory of Subjective Awareness

Joseph Levine, UMass Amherst

1. Introduction

First, a disclaimer. This paper is not intended to be a scholarly investigation of Sartre's theory of consciousness. Rather, I want to explore a certain problem in the treatment of conscious experience over the last several decades in the analytic tradition, and suggest that certain basic ideas that have their source, at least for me, in Sartre's account of consciousness in Being and Nothingness might point the way toward a solution. Let me begin by laying out the problem.

Perhaps the easiest way to introduce my topic is to ask the simple question: what distinguishes a conscious state, an experience, from an unconscious mental state? In Thomas Nagel's (1974) famous "bat paper" he gave the following answer: conscious states are those for which there is something it is like for the subject having them. He argued there, and elsewhere, that conscious experiences were essentially "subjective", involving a particular "point of view" on the world. It was this essential subjectivity of conscious experience, he believed, that made it so difficult to incorporate into the natural, physical world.

The "what-it's-like-for" formulation of conscious experience directly expresses the two aspects of conscious experience that have bedeviled attempts to naturalize it: phenomenal properties, or qualia, the features of experiences that determine what it's like for the subject; and subjectivity, an experience's being like something for the subject. It is the second aspect that is constitutively tied to the idea of a point of view, and is what I want to explore in this paper.

While Nagel argued that the subjectivity of conscious experience would make it extremely difficult to naturalize, many philosophers have responded with attempts to show that subjectivity is not the obstacle to naturalization that Nagel thought it was. But even before addressing the question of naturalizability, or reducibility, it's important to clarify just what it is for a mental state to be subjective in this way, to embody a point of view. Over the last several decades there have been many different theories of consciousness, and each one has had something to say about what subjectivity is. It seems to me that, when you abstract from many of the differences in detail, one can see two basic approaches: what I will call the "secondary awareness" approach and the "access" approach. Interestingly, this division cross-cuts the division between reductivists (materialists) and non-reductivists (who are often anti-materialists, though not always). What I will argue is that neither approach is successful in the end. What's needed is a third approach, the one that is inspired by aspects of Sartre's account in Being and Nothingness. I'll call it the "for-itself" approach.

2. Access Approach

Ned Block (1995) has famously distinguished between two notions of consciousness: "access consciousness" and "phenomenal consciousness". By calling a mental state "access conscious" Block means that the state in question is available to higher-level processes of reasoning, action planning, speech, and the like. "Phenomenal consciousness" is supposed to capture the Nagelian "what it's like" feature of conscious experiences, or what is also called an experience's "qualitative character". On Block's view, the phenomenal, or qualitative character of a mental state is an intrinsic feature, the one that presents such a problem for materialist reduction, the

source of the explanatory gap. Block insists that phenomenality is quite distinct from access, arguing that one can find cases, both hypothetical and actual, of mutual disassociation between the two.

While Block resists the idea that access is essential to any form of consciousness, many theorists see it as the key to understanding the subjectivity of conscious experience, what it is for a state to be for a subject. For example, Michael Tye's (1995 and 2000) "first-order representationalist" account of phenomenal consciousness, his so-called PANIC theory, make access an essential feature of the account. "PANIC" stands for Poised Abstract Non-conceptual Intentional Content. On the PANIC theory, what determines what it is like for a subject to have an experience is the intentional content of the mental representation that constitutes her having the experience. That the content in question is non-conceptual is meant to capture the apparent difference in format between paradigmatically phenomenal states like perceptual experiences and cognitive states like thoughts and beliefs, for which there is no phenomenal character (according to Tye). But since there could easily be unconscious representational states that have the same non-conceptual contents as their corresponding conscious states - as in subliminal perception - we still need to know what distinguishes conscious states from unconscious ones. This is the work that the "poised" in PANIC is supposed to do. When a state is poised to affect reasoning, deliberation, etc. in the right way - in other words, when it is access conscious - that is what makes it conscious. Subliminal perceptions are unconscious precisely because they are not poised to affect higher-level cognitive processes in the right way (though they do have noticeable effects, which is why psychologists believe they exist).

While first-order representationalist theories are explicitly reductive, it seems to me that there are non-reductive versions of access theory as well. For instance, among those who do not claim to reduce phenomenal character to anything else there is a debate concerning whether a state of conscious awareness must contain some reflexive conscious awareness as an essential component. In the phenomenological tradition, this idea of a secondary awareness is often associated with Brentano.[1] But a number of contemporary philosophers of mind who have no reductive ambitions still eschew the commitment to a reflexive component within a conscious experience. On their view, what makes a mental state phenomenally conscious is just its having this special mental feature, phenomenal character. While some, like Siewert (1998), argue that phenomenal character is essentially intentional - it presents how things look, sound, feel, etc. - they do not think any self-directed intentionality is an essential part.

Even though their analysis of phenomenal consciousness appeals to the looks and feels of experiences and not to any explicitly self-directed, or reflexive components, it seems to me that something of the sort enters the account nevertheless. Where does it come in? Well, it comes in by way of introspection. That is, one way of distinguishing conscious mental states from the unconscious ones is to say that the conscious ones are the ones to which we have introspective access, and this accessibility is essential to them. Unlike the reductive first-order representationalists, the non-reductive intentionalists do not attempt to provide a functional account of access. Still, the idea that a kind of subjective accessibility is essential to what it is to be a conscious state

[1] See especially Thomasson (2000), who explicitly rejects the reflexive view.

seems to be crucial to what distinguishes the conscious from the unconscious. It is the way such a view can capture the "for the subject" in "what it is like for the subject".

So what is wrong with the access approach? When it comes to the reductive versions there are, of course, all the explanatory gap problems associated with the reduction of qualitative character to either a functional or a physical property. But that isn't a problem with the access move itself, and it certainly doesn't apply to the non-reductive versions of the approach. Rather, it seems to me that the problem is that access - or, better, accessibility - is a dispositional feature, and subjectivity, what it is for a state to be <u>for</u> a subject, isn't adequately captured by a dispositional account of it.

I don't have a lot to say to support this claim, other than what I think is strong intuitive backing. As I look at my computer screen right now as I type these words it seems very clear to me that I embody a point of view on it, not just in the sense that I see it from a certain spacial perspective, but in the sense that the experience is <u>for</u> me; and its being <u>for</u> me is essential to what makes it a <u>conscious</u> experience.[2] But if this is right, than what it is that makes it <u>for</u> me must be something happening right now, in the moment, and thus an occurrent feature of the experience. That is, it isn't by virtue of the mere dispositional fact that I could, if I so desired, turn my conscious light on my own experience, that makes it conscious. Rather, it is by virtue of something it has right now that it is consciously experienced - as I keep repeating, it is <u>for</u> me. To my mind, as I

---

[2] I can hear Brie Gertler asking, "Is this a conceptual truth?" I don't know how to answer this. I certainly don't think of it as an analytic truth, just an analysis of what is meant by "conscious". Rather, it seems to me something revealed by reflection on our experience, to which we have first-person, intimate access. But what sort of truth is this? Is it straightforwardly empirical? If pushed I guess I'd say so, but I find the taxonomy of "empirical or conceptual" not quite adequate for philosophy. This is of course a large topic in itself.

just said, this means that the subjectivity of the experience is something occurrent, and not captured by any notion of accessibility, whether reductive or not.

3. Secondary Awareness Approach

Perhaps the most straightforward way to capture the idea of an occurrently subjective feature of conscious experience is to posit a secondary awareness as intimately tied to the primary awareness involved in any conscious experience. After all, isn't it my being aware of the experience what its being for me consists in? This basic idea motivates both reductive and non-reductive versions of the secondary awareness approach.

The reductive version of this approach is represented by the so-called "higher-order theory" of consciousness. David Rosenthal (1997), its most prominent proponent, presents the foundation for the theory this way. He says, expressing the point just made above, that conscious mental states are those mental states we are conscious of. For him this is a truism; it's just what we mean by calling a mental state "conscious". Certainly there is something deeply intuitive about this. If asked to say what distinguishes paradigm conscious mental states, such as the visual experiences I am now having of the computer screen in front of me, from paradigm unconscious mental states, such as the Freudian repressed desires I now battle, or the Chomskian states of my language device, a natural thing to say is that I am aware of, conscious of, the former, whereas I am unaware of, not conscious of, the latter. Rosenthal introduces the distinction between "transitive consciousness" and "intransitive consciousness", where the former is a matter of being conscious of some mental state, whereas the latter is a mental state's having the property of being conscious. He argues, on several grounds,

that the right way to approach a theory of consciousness is to take transitive consciousness as basic and intransitive consciousness as derivative. That is, a mental state is intransitively conscious just in case one is transitively conscious of it.

Rosenthal's particular version of higher-order theory is "higher-order thought" theory, which is distinguished from other versions like "higher-order perception" theory, as defended by Lycan (1997). The distinction turns on whether the awareness of one's mental state is in the form of a thought, a cognitive state, or something more like a perception. I think there are pros and cons to each of these choices, but for present purposes the difference won't matter. What both versions share is the idea that being aware of, conscious of, one's mental state is in what its being conscious consists.

Higher-order theory is unabashedly reductive. The first step in its reduction of consciousness to something else more tractable is to identify awareness, this transitive relation, with mental representation, as it is understood in cognitive science. Mental representations, on this view, are like other representations - natural language sentences, pictures, maps, etc. - in the sense that they are physically realized tokens that have both an intrinsic structure and a semantic relation to entities external to them. Just like the sentence "David Rosenthal is a prominent philosopher of mind" has intrinsic physical features and a meaning - it is about a particular man, and ascribes a property to him - so too mental representations have physical features - presumably neural properties - and also are about objects and their properties. While the analogy between the physical features that individuate sentences and those of mental representations seems straightforward enough, problems arise when comparing their semantic, or intentional features. The problem is that if we ask what it is about a

sentence of English, like the one above, by virtue of which it is about a man and a property being ascribed to him, the most plausible answer is that it derives ultimately from the mental states of users of the language. It seems uncontroversial that if humans had never lived but the wind caused a shape to form in the sand on a beach of the same form as our sentence above, it wouldn't be about anything; it would be meaningless. But on pain of infinite regress (or appeal to God - and then, what makes His states meaningful?) something similar can't be said about our mental representations.

This problem, that the intentional "buck", as it were, stops at our minds, with their mental representations, has given rise to the "naturalizing intentionality" industry.[3] It is notorious that after all the ink that has been spilled on this topic that no consensus has been reached on how to naturalize intentionality. To my mind, the most promising approach is the "informational", or "nomic-covariation" approach. On this view, a physical token of a certain type in the brain is about an object/property by virtue of a causal law that correlates instances of the object/property in question with tokens of the type in question. Of course lots of bells and whistles need to be added to this basic formulation in order to overcome problems about the possibility of misrepresentation, and other, related problems. As far as I know, no account of these bells and whistles has been shown to be counter-example free. Nevertheless, if a genuine reduction of conscious awareness to mental representation, as envisioned by higher-order theory, is to occur, some such account of intentionality must be found.

---

[3] Some of the seminal work in this area is Dretske (1995), Fodor (1990), and Millikan (1989).

I will set aside the problem of naturalizing intentionality and just assume some solution along the lines of informational semantics can be found. The fundamental idea, then, is that the intentionality of a mental state derives from the intentionality, or semantics, of the relevant mental representation. Therefore, on higher-order theory, the consciousness of a mental state has to be cashed out as the mental representation of that state (together with certain added conditions on the nature of that representation). The dispute between higher-order thought theory and higher-order perception theory is thus very much about the format of the relevant higher-order representation.

One of the added conditions on the higher-order mental representation that Rosenthal emphasizes is that it have the form of a self-ascription: that it say something like "I am now perceiving a computer screen", rather than merely "there is a perception of a computer screen going on in such-and-such a location". One reason for insisting on the self-ascriptive format is precisely to capture the essential subjectivity of conscious experience, that it involves there being something it is like for the subject. Representing oneself (with a mental "essential indexical") as undergoing a certain perceptual state, say, is what captures that idea of being "for the subject".

We now have the principal elements of higher-order theory before us, or at least those that will concern the topic of this paper. So what are the problems? To begin with, a standard objection to higher-order theory is that it leads to a regress problem. If a first-order state, such as a perceptual state, is conscious by virtue of another state, the higher-order one, constituting consciousness of it, then the higher-order state must be conscious as well. But now we need to say what makes the higher-order state conscious. Clearly, appealing at this point to an even higher-order state leads to an

infinite regress of states. As stated, however, this isn't a real problem for higher-order theory, since its advocates are quite clear that they do not claim that the higher-order state by virtue of which the lower-order state is conscious is itself conscious. Quite the contrary. It is by virtue of the higher-order state's representing oneself as occupying the lower-order state that the latter is conscious. It's crucial for the view that one is not conscious of the higher-order state. In fact, introspection alone reveals that; after all, as I currently am enjoying a conscious visual experience of the computer screen in front of me I am not aware of any higher-order thought or perception. If the higher-order state had to be conscious this would make the theory quite implausible from a phenomenological point of view, in addition to leading to an infinite regress.

There are two counter-replies that critics of higher-order theory make at this point. First, some wonder how it is that by having one unconscious mental state represent oneself as being in another (previously) unconscious state that one comes to be in a conscious state. Put another way, how can an unconscious mental representation make an unconscious state into a conscious one? To this challenge higher-order theorists respond that the objection reveals a serious misunderstanding of the view. It isn't that the higher-order state makes the lower-order state conscious, as if this were a causal process. That indeed would be mysterious. Rather, the two states, standing in the relevant representation relation, together constitute the conscious experience in question. This is a reductive theory, so one can't expect to find consciousness itself appealed to in the explanans. On the contrary, argues Rosenthal, it's only by analytically breaking the allegedly monadic property of being-conscious into these non-conscious components - a representation relation holding between two

unconscious states - that can one even hope to explain consciousness. As Fodor famously quipped about intentionality - "if it's real, then it must really be something else" - so too for consciousness.

The second reply is similar in spirit, but doesn't confuse the constitutive nature of the higher-order account with a causal one. Siewert (2013) puts the point this way. The original intuition - indeed, alleged truism - that supported higher-order theory was the principle that conscious states are those mental states we are conscious of. This is what motivates the principal idea that intransitive consciousness must be analyzed in terms of transitive consciousness. But notice that in this formulation the phrase "conscious of" appears, not "represented by". If we're to honor the alleged truism, then our relation to the target state must be one of being conscious of it, and thus the higher-order advocate is faced with a dilemma: either honor the fundamental principle on which the theory is based, but then face the regress problem, or avoid the regress problem, but lose the support claimed from the fundamental principle, the alleged truism. The first horn is clearly unacceptable. While the second horn does not lead to any outright refutation, it does, so the argument goes, undermine that initial intuitive support claimed by advocates of the theory.

A plausible rejoinder to this argument is basically the same as to the previous one: higher-order theory is attempting to explain consciousness, which clearly cannot be done by taking it as primitive. Therefore, no analysis of what it is for a mental state

to be conscious can rest with the idea that it is a state we are, literally, conscious of.[4] The point of appealing to the truism is just to show that what many theories take to be a monadic property of mental states is really a relation between the subject and the state in question. It is just another way of capturing the "for the subject" in the characterization of conscious experience in terms of "what it is like for the subject".

Here is one way to see what's at stake here. Two opposing reductivist programs are higher-order theory and first-order representationalism. On the latter view, conscious states are first-order representational states - paradigmatic cases are perceptual states - that meet certain further conditions. While Rosenthal, the standard bearer for higher-order theory, battles under the slogan "conscious states are states we are conscious of", Dretske (1997), one of the standard bearers for first-order representationalism, captures the view with the slogan, "conscious states are those we are conscious with". Neither Rosenthal nor Dretske intends to leave the "conscious" in "conscious of/with" as a primitive. Both want to analyze it as a matter of mental representation, a notion that too requires some analysis, but one that seems much closer to having a realization theory than an unreduced notion of consciousness does. So in their dispute what matters most is the preposition following the term "conscious", not the term "conscious" itself.

As I stated at the start of the paper, I don't think higher-order theory does adequately capture subjectivity in the end. Part of the problem is that I don't think any

---

[4] Siewert supports his criticism by appeal to a principle he calls "IC" (for "is conscious") that reads as follows: "Every state of being conscious of something is a conscious state." (Siewert 2013, page 244) He argues that IC is every bit as intuitive as CO (for "conscious-of"), Siewert's name for the principle appealed to by Rosenthal to the effect that a conscious state is one we are conscious of. As I see it, IC is sustainable, even by the higher-order theorist, so long as we modify CO so that the "conscious of" isn't taken literally, for reasons argued in the text. That way, we can still maintain, in keeping with IC, that if we are literally conscious of something, our being so involves our being in a conscious state.

reductivist account of conscious experience works, neither for subjectivity nor for qualitative character. But the objection I want to press here against higher-order theory isn't a general worry about reduction; it's quite specific to the way it tries to reconstruct subjectivity, a state's being for the subject. I want to introduce the problem I see by drawing an analogy with another issue, one that exercised philosophers some decades ago when appeal to mental representation became popular in the context of cognitive science.

When it first became fashionable to appeal to mental representations (again) to explain behavior, a challenge that was often pressed against this sort of explanation went like this: Genuine representations are representations for someone. To represent X is to represent X to or for some subject who utilizes the representation in some way. But, went the objection, this would entail positing some sort of homunculus to interpret and use the representation, to be the subject for whom the representation represents what it does.[5]

The standard computationalist response to this objection was to say that mental representations are, at least ultimately[6], self-interpreting, or self-understanding, so one doesn't need a subject for whom it represents and who must count as understanding the representation. But how do they then cash out the "self" in "self-understanding"? Basically they functionalize it. Mental representations are for the subject possessing them when they play the appropriate role in the cognitive economy that constitutes the subject. So in a sense they are happening in the subject, they are not something the

---

[5] A version of this objection can be seen in Ryle's (1949) attack on the "intellectualist legend".

[6] I say "ultimately" because there may indeed be homunculi of a sort for certain cognitive systems communicating with others.

subject explicitly takes an attitude toward, except in the somewhat deflationary sense in which playing the functional role in question realizes the relation of taking an attitude toward the relevant representations.

With respect to the project of legitimating appeal to mental representation in cognitive science I think this response to the challenge is fully adequate. But notice what this does to the higher-order theory. We start by asking: what is it that makes a perceptual state conscious? We then immediately analyze what it is for a state to be conscious in terms of its subjectivity, its being like something for the subject. But of course we still need an account of this notion itself, and higher-order theory responds by positing an extra representation, one over and above the state to which we want to ascribe consciousness. But the question can still be asked: what makes this representation itself, the higher-order one, for the subject? The answer, in the end, is just the same as the one we found for the access approach: it's a matter of its functional role, and thus its dispositional profile.

My point is that insofar as we are trying to capture the idea of an occurrent for the subject, higher-order theory doesn't do much better than first-order theory. The reason is that we still primarily get the effect of the "for-ness" through functional role, which makes this a kind of access approach. If we were unhappy with the access account in the first place because it doesn't adequately reconstruct the relation of being for the subject, making it dispositional, rather than occurrent, my claim is that the same problem attends the higher-order theory.[7]

[7] So, in a sense, I'm adopting a version of Siewert's regress argument. That is, rather than IC, I have IF: Any state by which an object is "for the subject" is itself "for the subject".

One reaction to the criticism I've mounted of higher-order theory is to say that the fundamental problem stems from its reductive ambitions. Indeed, I am quite sympathetic to the idea that no reductive theory can reconstruct the full-blooded notion of subjectivity. (Reductivists would agree, on the whole, though they'd claim that that's because there is something either incoherent, or just fictional, about the full-blooded notion. Their claim is that their theories can capture what in reality there is to capture. But this isn't the place to play out that battle.) So let's turn now to a non-reductive version of the secondary awareness approach, and see whether it can properly account for the subjectivity of experience.

The kind of approach I have in mind here has been attributed to Brentano (see Thomasson 2000, for instance) in a number of recent discussions of consciousness. The idea is that each conscious experience contains within itself a primary intentional object - whatever it is we are consciously aware of at the time - and a secondary intentional object, the conscious experience itself. So, for example, suppose I am visually perceiving a ripe tomato sitting on my kitchen counter. Let's assume there is nothing funny going on and the experience is fully conscious, though I am not especially reflecting on it, or engaging in any kind of introspection. I simply see it, fully aware of its presence and its properties (such as being round and red).

On the theory under consideration, my conscious experience contains two objects: the ripe tomato on the counter and my visual experience of the ripe tomato itself. However, the two objects do not have equal billing in my current theater of consciousness. The tomato is the primary object, whereas the experience itself is merely the secondary object. But how should we understand this distinction? What

work does it do? Also, how do we understand the two awareness relations posited here, the awareness of the tomato and the awareness of the experience? Are these relations substantially of the same kind, just with different objects - indeed, with objects of different grades - or are the relations quite different in character, which would explain the difference in the kinds of objects they take (one taking primary objects and the other secondary ones)?

Let's begin with the relation. In common with its reductive sibling, higher-order theory, the non-reductive secondary awareness approach begins from the alleged truism that conscious states are states we are conscious of, or aware of. If this is right - if we're going to reconstruct being <u>for</u> the subject in terms of awareness of the state by the subject - then it's hard to see how the awareness involved in secondary awareness can be different in kind from the awareness involved in primary awareness. This echoes Siewert's earlier challenge to higher-order theory, though now there need be no explicit commitment to the effect that secondary awareness isn't itself a kind of conscious awareness.

However, if the awareness in secondary awareness really is of the same sort as that involved in primary awareness, then we are faced immediately with the regress problem again. After all, there is no attempt here to reduce the consciousness of a single state to a naturalistic representation relation between two non-conscious states, as with higher-order theory. Thus we can't escape the requirement that the secondary awareness be itself susceptible to the analysis that attends the original conscious experience as a whole. If so, then it seems in order for the secondary awareness to be

a kind of conscious awareness that it must be attended by a “tertiary” awareness of it, and then we’re off and running into the depths of infinite regress.

One way of making the primary-secondary distinction actually reinforces the regress problem. If asked, in what does the secondary awareness’s being secondary consist, a natural answer appeals to differences in attentional focus across the objects of consciousness. When I’m absorbed with seeing the tomato - say, as I stare at it while deciding whether or not to include it in the salad I’m making - my attention is fixed on the tomato, not on my own experience. Just as we know that visual attention selects out certain among the items visually perceived for increased processing, yet without totally losing the perception of objects not within the current scope of attention, so too, on this view, one maintains the awareness of one’s own experience while attention is focused on its primary object.

On this model what distinguishes primary from secondary awareness is whether or not the object of awareness lies within the focal area of attention. But this really makes the two forms of awareness seem very much alike, especially since, as one might expect from such a model, it is possible to switch attentional focus so that the secondary object becomes primary; this is one model of introspection. But if this is the case, if the awareness relation itself is pretty much the same for both primary and secondary awareness, then it does seem that the regress problem becomes all the more pressing.

While I think there really is a regress threat against the secondary awareness view, there are ways to counter it. For one thing, Brentano’s view is often described as a “self-representational” view, though the “representational” part is a modern, reductive

add-on. The idea is to beat back any regress problems by treating a conscious state as not a combination of two conscious states, but rather as a single state that has two intentional contents: the primary object and the state itself. Given there is only one state involved, which itself possesses all the conditions (including "self-representation") necessary for being a conscious state, there is no way to get the dreaded regress going.[8]

Now, as developed in a reductive manner by Kriegel and others, I don't think this move works. As I have argued elsewhere, the kind of mereological view that underlies reductive self-representational accounts can't really support a substantive distinction between higher-order theory and so-called "same-order" theory. So long as the self-representation function is reserved for a constituent of the entire state, I don't see that one can lean much on the alleged unity of the entire state to avoid the problems that attend the explicitly two-state higher-order theory. But these consequences largely follow from demands brought on by the reductive enterprise of analyzing awareness as representation, naturalistically construed. If we leave awareness unanalyzed, then it may be easier to avoid the regress problem.

The idea is this. My conscious visual experience of the ripe tomato indeed constitutes a single mental state. On this non-reductive version of the view, the state itself does not decompose into parts, each of which stands in its own awareness relation to its object - whether primary or secondary, a move required by the reductive analysis in terms of physically realized mental representations. Rather, this one state

---

[8] See Kriegel (2009), and also Gennaro (1996) for a defense of a version of self-representationalism he calls "WIV", for "wide intrinsicality view", and also Gennaro (2002), in which he attributes an analogous view to Sartre. More on this below.

involves the subject's simultaneously standing in two awareness relations, one to the primary object and the other to the state itself.

I think it's just possible this view can avoid regress, but only by taking on board some very implausible features. The basic problem is this. Regress gets off the ground once one analyzes consciousness in terms of the conscious-of relation. If you want to say what makes my visual experience of the ripe tomato conscious is that I am conscious of it, then one faces a dilemma: either by "conscious of" one is using "conscious" to mean the same thing one means in "is conscious" or one isn't. If they do mean the same[9], then it's hard to see how to avoid the regress. If being aware of a state is what it takes for it to be conscious, then you have to say the same thing about the consciousness of the state itself. If, on the other hand, one doesn't mean the same thing, then one is engaged in a reductive project after all, though it needn't be reduction to a naturalistic, or physically realizable relation.

To elaborate on this second horn of the dilemma, consider again our example of my visual experience of the ripe tomato. This experience is a single conscious state, embodying two instances of the awareness relation. However, whereas we can freely characterize my overall visual experience of the tomato as my being conscious of the tomato, when we now break down that consciousness of the tomato into its component relations, primary awareness of the tomato and secondary awareness of the state of being primarily aware of the tomato, we cannot appropriately characterize either of these awareness relations as "conscious of" relations. There has to be something about the awareness that figures in my basic relation (i.e. basic with respect to the non-

[9] As in Siewert's (2013) IC principle. See footnote 4 above.

basic consciousness relation) to both the tomato and to my experience that doesn't quite qualify as full-blooded consciousness.

The non-naturalistic yet reductive theory would then look like this. Consciousness is a relation one stands in with respect to an object when one stands in a more basic awareness relation to the object and simultaneously to the state of being aware of the object. But what is this awareness relation? It's not the representation relation holding between a token mental representation and its object. It's not what we normally identify as consciousness. Is there such a thing as awareness of this kind? I find it mysterious and phenomenologically implausible. To me, it's precisely my awareness of the tomato that bears all the marks of consciousness. I find I can't discern a kind of awareness that isn't consciousness itself. This isn't a knock-down objection, but it does make the view harder to accept.

To my mind, when I have a visual experience of an object - say, the ripe tomato - it is that very awareness of the tomato that is the instance of consciousness at issue here, and this includes the feature of subjectivity. That is, I want to say that the awareness of the tomato is a form of conscious awareness (at least in part) by virtue of the fact that it - that very awareness of the tomato, not some other awareness stuck on top of it - is a subjective taking of a point of view on the tomato, which, to me, is another way of saying that my awareness of the tomato is for me. In other words, I don't know what awareness is if it isn't full-blooded conscious awareness, and taking this line seems to preclude the secondary awareness approach to analyzing consciousness.

4. For-Itself Approach

As I stated at the beginning of the paper, this is not intended to be a scholarly interpretation of Sartre's theory. Furthermore, there are elements of (what I take to be) his view that I do not endorse and will not incorporate into the approach I want to outline here. Still, I think enough of what I have to say is legitimately to be found in Sartre's work that it is appropriate to call this a Sartrean approach - or, as in the title, a "quasi-Sartrean" approach - to the question of subjectivity.

The fundamental distinction in the Sartrean framework that defines the For-Itself approach is that between the "in-itself" and the "for-itself" (or, between "being" and "nothingness"). The subjectivity of a conscious experience, its constituting a point of view on its content, or there being something it's like <u>for</u> the subject, is a matter of the holding of a primitive consciousness relation between the subject and the object. The subjectivity of the relation is an essential feature of the relation, not a constituent of the relation. That is the crux of the approach. Let me now elaborate on the distinctions involved (including between being a feature and being a constituent), show how they are brought together to explain subjectivity, and then delineate how this approach both incorporates features of the other two approaches while avoiding their problems. (Of course this doesn't mean it has no problems of its own.)

For Sartre, as I understand him, the distinction between the in-itself and the for-itself is intended to mark a fundamental ontological cleavage in reality. I do not endorse all of the features of this distinction that play crucial roles in Sartre's philosophy (especially his account of human freedom), but I do want to take on board the idea that the for-itself, consciousness, is a basic phenomenon in nature, not reducible to the in-itself. So what is the distinction, as I employ it? There are two features of the in-itself

that distinguish it from the for-itself: the in-itself encompasses all of the objects and properties that constitute concrete, material reality - the universe, with all that composes it, and all that is subject to its fundamental and derived laws of nature; and also, it provides the contents for the conscious awareness that is the for-itself.[10] The for-itself, on the other hand, is just the pure relation of conscious awareness itself, with no substantial reality in its own right, but ontologically exhausted by its content - what it is aware of - and the relation of being aware of it.

The insubstantiality of the for-itself, the conscious subject, is what suggests the appellation of "nothingness". Whether there is a more fundamental ontological base underlying the structural relations between the for-itself and the in-itself, one that constitutes both the subject and object of a conscious experience, is not an issue on which I currently take a position; clearly Sartre thought there was not. What matters for me, however, is the fact that this pure relational, point-of-view-ness of the for-itself is all that is revealed within the (Cartesian?) theater of conscious experience.

The idea that consciousness is fundamentally a relation, and a fundamental relation at that, explains a number of features that many philosophers, of various persuasions, have emphasized. So, for instance, Hume (2012, Book 1, Part 4, section 6) famously denied that we have any impression of the self. Mental states, according to Hume, consisted of the having of impressions and ideas, faint copies of impressions. Impressions are always of something, whether it be colors, shapes, sounds, or bodily sensations. In each of these cases there is a content that the impression delivers to the

---

[10] Actually, on my own view of the consciousness relation, which I dub the "AA relation" (for Acquaintance/Appearance), the objects of consciousness are not found within the realm of objective concrete reality, but rather have the status of "virtual objects" (see Levine 2010). But this difference won't matter for present purposes.

conscious mind, and the consciousness of this content exhausts the mental state in question. Since ideas - the Humean stand-in for concepts, the material of thought - derive their contents from the impressions of which they are copies, there would need to be an impression of the self if we were to have a substantial idea of the self. But try mightily as he did, Hume, as he surveyed the contents of his conscious experience, could find no impression of self. Well, if consciousness, the for-itself, is a kind of "nothingness", then this isn't surprising. In other words, if all there is to the subject of experience is its being the point of view on the contents of the experience, and nothing more, then there couldn't be an impression of it, as there really isn't anything there to constitute the content of the impression.

As I said above, though for Sartre the ontologically special status of the for-itself as nothingness is clearly of prime importance, my point doesn't rely on this. What matters is not so much whether or not there actually is something in the world (e.g. the brain), in the realm of the in-itself, that can be identified with the subject of experience. Maybe there is; indeed, it seems likely. But two points need emphasis. First, I doubt whatever object we might find to count as the subject will support the heavy-duty intuitions of personal identity that we seem to have. Second, more important for our discussion, within the relation of conscious awareness the subject does not figure as content, as element of which we are aware, but rather as the implicit audience member, as it were, for the point of view <u>on</u> the contents that constitute the experience.

A point related to Hume's is expressed in contemporary discussions of the so-called "transparency" of experience. The idea that phenomenally conscious states are transparent is principally associated with reductive representationalists. Following

Harman (1990), they argue that the allegedly queer features of phenomenal experience - the phenomenal properties, or qualia - are in fact the everyday features of the external objects we experience. Where anti-materialists claim to find the queer property of a reddish qualitative character as a feature of our visual experience, Harman and his followers claim to find only the redness we see in the ripe tomato. They argue that if you reflect clearly on your experience, you'll see that you can find no special features of the experience over and above the features of the objects that your experience presents to you. Thus phenomenal, or qualitative character, on this view, is identified with the intentional content of the relevant perceptual representation. Once you add a naturalized theory of mental representation, you have the ingredients for a fully reductive theory of conscious experience.

While I do not buy the reductive theory being sold here, I have always thought there is something right about the transparency intuition itself. For one thing, it supports the idea that conscious experiences are essentially intentional - that phenomenal character is a kind of intentional content, and that much I do buy. But more than that, what seems right here is that it isn't one's state of awareness itself that one is aware of, or any special features of it, but rather the objects of which one is directly conscious - that ripe tomato on the kitchen counter. The idea is that there is nothing that can be made an object of awareness within a state of awareness over and above the original object of awareness itself. Nothing to see her but the tomato and its features.

Interestingly, I think the Sartrean theory of the for-itself does a better job of explaining transparency than does reductive representationalism. After all, if the visual experience of the ripe tomato is constituted by the physical tokening of a mental

representation, then there is no reason in principle why introspection couldn't detect features of the representation itself, over and above its content; and therefore no reason one couldn't identify the phenomenal redness one experiences when looking at the ripe tomato (as opposed to the redness of the tomato itself) with features of the representation itself - the "mental paint" dismissed by Harman and others. But if conscious awareness is insubstantial, nothing more than a point of view, a pure awareness, of whatever is its object, its content, then the metaphor of "seeing through" the experience to what the experience is of seems perfectly natural, indeed the only way it could be.

There are two features of the for-itself that are crucial for our discussion: first, that it is insubstantial, and therefore ineligible for being an object of awareness itself, constituting a pure point of view on genuine objects, in-themselves; and second, that it is in fact for itself. By this I mean that the awareness relation that constitutes the for-itself is intentionally directed on its object, but inherent in this form of intentional directedness is the awareness's being for the subject. But what does for-ness come to for Sartre?

Gennaro (2002) argues that subjectivity for Sartre is captured by a kind of secondary awareness. There are certainly many passages that Gennaro quotes that seem to support his interpretation. For instance, the following seems especially clear on this point:

"the necessary and sufficient condition for a knowing consciousness to be knowledge of its object is that it be consciousness of itself as being that knowledge. That is a necessary condition, for if my consciousness were not consciousness of being

consciousness of the table, it would then be consciousness of the table without consciousness of being so. In other words, it would be consciousness ignorant of itself, an unconscious - which is absurd." (Sartre 1956, page 11)

While I grant the import of this quote (and others Gennaro cites), it seems to me that there is a way to capture the subjectivity Sartre is after here, using his idea of the primitive relation that the for-itself brings to the world, without positing a secondary awareness. Let me explain.

First let me introduce another couple of Sartrean distinctions of relevance for this discussion: that between "pre-reflective consciousness" and "reflective consciousness" and between "positional" and "non-positional" consciousness. Pre-reflective consciousness is essentially the outer-directed form of awareness we've been discussing, as when I visually experience that ripe tomato on the kitchen counter. Reflective consciousness, on the other hand, is achieved when one trains one's awareness on oneself, or one's experience, thus a form of introspection. Positional consciousness is a matter of the focal object of one's awareness, and involves "positing" an object as the content of awareness. "Non-positional" consciousness refers to what surrounds, as it were, one's focal awareness, and does not involve treating its contents as objects. So, Sartre claims at one point, speaking of pre-reflective consciousness, "every positional consciousness of an object is at the same time a non-positional consciousness of itself" (Sartre 1956, page 13). This too provides support for Gennaro's interpretation.

Clearly what exercises Sartre in the first passage quoted is that a pre-reflective conscious experience that isn't somehow "taken up" by, or of significance to - that is, is for - the subject, just isn't a conscious experience. But must one posit a secondary awareness relation in order to capture this? I don't think so, unless, of course, one takes one's analysis of conscious experience to be part of a reductive project. If that's the case, then I think there is good reason to go that route. That is, if one only has natural causal relations out of which to construct the consciousness relation, then perhaps the only way to capture subjectivity is via an analysis of consciousness into constituent representation relations, one of which is outer directed and the other inner directed. This is how Gennaro does it (and is the basis of higher-order theory). But, I now want to argue, if you take the consciousness relation as a primitive, not analyzable or constructible out of naturalized representation relations, then another option presents itself.

On the secondary awareness approach, whether of the reductive or non-reductive variety, the overall conscious experience is decomposed into two constituent relations: awareness of the object, and awareness of itself. When these two relations hold, then one's state constitutes a conscious experience. On the reductive approach, given that awareness is cashed out as physically realized mental representation, the only way to get subjectivity into the state is to posit this secondary awareness - a second mental representation. There is no way to literally build the subjectivity into the primary awareness relation, given its status as an unconscious mental representation.

So you build it into the conscious state by treating the overall conscious state as a mereological sum of two constituent representation relations.[11]

But if one is a non-reductivist, then there is another way to build subjectivity into the primary, "pre-reflective" awareness: make it a feature of the primary awareness relation itself, rather than a constituent, or component awareness relation. That is, for-ness isn't a matter of a second awareness accompanying the primary one, but rather being for the subject is part of what it is to be aware of something. Awareness that isn't for anything (or anyone) just isn't awareness.

Earlier I compared the question of what makes awareness for a subject to the question facing cognitive science of what makes representation for the subject in whom the token representation resides. I said that what legitimates appeal to mental representation in cognitive science is treating representation-for as constituted by the mental token's realizing the relevant functional role(s). I also argued that the functional account of being-for didn't seem appropriate for the way in which conscious experience is for the subject. As we saw with higher-order theory, merely adding on another layer of awareness - or representation - doesn't solve the problem since the question of what makes that added layer for the subject just recurs. As far as I can see the for-ness of experience is not something that can be analyzed. Rather, it's just inherently part of what it is to be a conscious awareness of something that that awareness be for the

---

[11] As Gennaro (2002, page 307) puts it: "What else is a 'non-positional self-awareness' except some kind of non-conscious meta-psychological mental state? I cannot understand it any other way." I think within the reductive program this is exactly right.

subject. Again, that's all there is, on this view, to being a subject: being that for which/ whom the awareness exists, for whom it has significance. Hence the appropriateness of saying that the conscious subject is entirely captured by calling it a "for-itself".

Above I objected to the non-reductive secondary awareness approach that it had even more of a regress problem than did its reductive cousin. Sartre, in passages like the following, clearly worried about this:

"Consciousness of self is not dual. If we wish to avoid an infinite regress, there must be an immediate, non-cognitive relation of the self to itself." (Sartre 1956, page 12)

Gennaro interprets Sartre here as blocking the regress by appealing to a kind of secondary awareness that is joined within the same conscious state as the primary awareness of the external object. This is the constituent analysis. The "non-cognitive" is understood to mean "non-positional", not positing the self as an object when the primary object of one's "positional" consciousness is something external.

But I think one gets a better position if one takes Sartre to be hinting at the view of subjectivity as a feature of the original consciousness relation to the object, this feature of being for-the-subject as a necessary condition of a cognitive relation constituting a state of consciousness at all.[12] Now, there is clearly something awareness-like about the notion of for-ness. The idea is most easily characterized as being of significance, being somehow "taken up", by the subject, as mentioned above,

---

[12] Some support for this reading of Sartre's intentions comes from this passage: "This [non-positional] self-consciousness we ought to consider not as a new consciousness...[it is] a quality of the positional consciousness." (Sartre 1956, page 14) Gennaro quotes this passage to support his self-representational interpretation, taking the point to be about there being a single state of which this non-positional self-consciousness is an integral part. But one can read the "quality" here in the feature, or adverbial sense I am advocating.

and this certainly is intuitively related to being aware of it. Given that conscious awareness is intrinsically subjective, we can't really aptly characterize the subjectivity as itself a kind of awareness, or we start down the path of regress, as we saw earlier. What's more, this means that awareness doesn't take itself as its own object, but rather counts as awareness of its proper object precisely because it makes that object (not the awareness of it) as of significance for the subject. One can think of this on the model of adverbial approaches. Subjectivity is not a separate element within awareness, but rather a way of being aware of an object. Only being subjectively aware is a kind of being consciously aware.

One might object to this alternative to the secondary awareness approach as follows.[13] When we do introspect - engage in what Sartre calls "reflective consciousness" - we find that there is a kind of epistemic immediacy to our encounter with our own conscious experience that is distinctive of the first-person perspective. If the original, pre-reflective conscious state didn't itself contain awareness of itself, how would the explicitly reflective, introspective attitude be capable of, as it were, "finding" the experience? How is it we so easily and immediately know what we are experiencing if it isn't by virtue of some self-awareness already embodied in pre-reflective conscious experience?

Indeed, Sartre himself explicitly links the non-positional consciousness of self with the ability to positionally reflect on one's conscious experience when he says, "it is the non-reflective consciousness which renders the reflection possible..." (Sartre 1956, page 13). I must admit that I don't have a worked-out view of introspection, so I can

[13] I am grateful to Tom McClelland for pressing a version of this objection.

only speculate here. What I'm inclined to say is this. Surely, if there is some kind of epistemic distinctiveness to introspection, it must have something to do with the epistemic immediacy, or acquaintance, embodied in first-order, pre-reflective conscious experience itself. But it isn't clear to me that whatever it is about pre-reflective conscious experience that enables taking a first-person perspective on it - positionally, as it were - requires that it already contain a form of secondary awareness within it as a constituent relation. It could very well be that the for-ness feature of the pre-reflective consciousness relation plays the requisite role. At any rate, I don't see why it couldn't.

5. Conclusion

Conscious experiences are not merely instantiations of phenomenal properties, but rather inherently intentional and subjective relations of a particular, ontologically fundamental kind. Philosophers have wrestled with integrating the subjectivity of the experience with its content, and traditional views tend to take either the secondary awareness approach or the access approach. I have argued that neither of these approaches achieves the kind of integration necessary to capture the subjective nature of a conscious experience of a typical object like a ripe tomato. Rather, the only way to integrate the two is to do just that - make them inseparable - as is found in the "quasi-Sartrean" approach presented here.[14]

---

[14] An earlier version of this paper was presented to a workshop on The Subjective Structure of Consciousness, at the University of Manchester, June 22, 2014. I would like to thank the participants in that workshop for their very helpful comments during discussion. In particular, I want to thank Aaron Henry, Tom McClelland, Donnchadh O'Conaill, and Sebastian Wazl for discussion on the topic of this paper.

REFERENCES

Block, N. (1995). “On A Confusion About a Function of Consciousness”, Behavioral and Brain Sciences, 18:2.

Dretske, F. (1995). Naturalizing the Mind. Cambridge, MA: Bradford Books/The MIT Press.

.....(1997). "Conscious Experience" in The Nature of Consciousness, Edited by Block. N., Flanagan. O., Guzeldere. G., MIT, Bradford.

Fodor, J.A. (1990). A Theory of Content and Other Essays. Cambridge, MA: Bradford Books/The MIT Press.

Gennaro, R. (1996). Consciousness and Self-Consciousness: A Defense of the Higher-Order Thought Theory of Consciousness, Amsterdam: John Benjamins Publisher.

..... (2002). “Jean-Paul Sartre and the HOT Theory of Consciousness”, Canadian Journal of Philosophy, 32:3, 293-330.

Harman, G. (1990). “The Intrinsic Quality of Experience”, in J. Tomberlin, ed., Philosophical Perspectives, 4, Action Theory and Philosophy of Mind, Atascadero, California: Ridgeview Publishing Co., 31-52.

Hume, D. (2012). A Treatise of Human Nature. Mineola, NY: Dover; originally published in London, 1739-1740.

Kriegel, U. (2009). Subjective Consciousness: A Self-Representational Theory, Oxford University Press.

Millikan, R. (1989). “Biosemantics”, The Journal of Philosophy, 86:6, 281-297.

Levine, J. (2010). “Phenomenal Experience: A Cartesian Theater Revival”, Philosophical Issues, 20, Philosophy of MInd.

Lycan, W.G. (1997). “Consciousness as Internal Monitoring”, in The Nature of Consciousness: Philosophical Debates, eds. N. Block, O. Flanagan, and G. Guzeldere. Cambridge: MIT Press.

Nagel, T. (1974). "What Is It Like to Be a Bat?", The Philosophical Review, vol. 82, 435 - 450.

Rosenthal, D. (1997). “A Theory of Consciousness”, in Block, N., Flanagan, O., and Güzeldere, G. eds., The Nature of Consciousness: Philosophical Debates, Cambridge: MIT Press, 729-753.

Ryle, G. (1949). The Concept of Mind, Hutchinson.

Sartre, J-P (1956). Being and Nothingness, Hazel E. Barnes, translator, New York: Philosophical Library.

Siewert, C. (1998). The Significance of Consciousness. Princeton, NJ: Princeton University Press.

Siewert, C. (2013). "Phenomenality and Self-Consciousness", in Kriegel, U., ed., Phenomenal Intentionality, Oxford University Press.

Thomasson (2000). "After Brentano: A One-Level Theory of Consciousness", European Journal of Philosophy, 8:2, 190-209.

Tye, M. (1995). Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind. Cambridge, MA: Bradford Books/M.I.T. Press.

Tye, M. (2000). Consciousness, Color, and Content. Cambridge, MA: Bradford Books/ M.I.T. Press.