

Précis of *Subjective consciousness: a self-representational theory*

Uriah Kriegel

© Springer Science+Business Media B.V. 2011

According to self-representationalism, a mental state is phenomenally conscious iff it represents itself in the right way. In *Subjective Consciousness: A Self-Representational Theory* (henceforth, SC), I develop and defend my own version of self-representationalism. The view, and the case for it, can be profitably divided in two: an account of the explanandum and an account of the explanans.

The explanandum, phenomenal consciousness, is often fixed upon with ‘what it is like’ talk. When I look at the blue sky, and have a conscious experience thereof, there is a bluish way it is like for me to have the experience. This ‘bluish way it is like for me’ is the experience’s phenomenal character. On my view, however, there is more to be said about phenomenal character—there is more structure to it than is typically recognized. In particular, I distinguish two components of the ‘bluish way it is like for me’ to have the experience: the bluish component, which I call *qualitative character*, and the for-me component, which I call *subjective character*.

To make a conceptual separation between qualitative and subjective character is not to imply that they can occur apart from one another. My view is that there are many determinate phenomenal characters—bluish-for-me-ness, greenish-for-me-ness, bitterish-for-me-ness, trumpet-for-me-ness, etc.—and the determinable of all of them is for-me-ness as such. We grasp what subjective character is by fixing on what is common to all phenomenally conscious states, and grasp what qualitative character is by fixing on what varies among them. In this way, we grasp the difference between qualitative and subjective character without requiring them to come apart. An upshot is that for any phenomenally conscious state, qualitative character is what makes it the phenomenally conscious state it is and subjective character is what makes it a phenomenally consciousness at all. In other words, qualitative character captures the *identity* condition of conscious states, while

U. Kriegel (✉)
Department of Philosophy, University of Arizona, Tucson, AZ 85721, USA
e-mail: theuriah@gmail.com

subjective character captures their *existence* condition. The latter is thus phenomenal consciousness *as such*.

For this reason, I regard subjective character as the central explanandum of the theory of consciousness. What we need to explain is phenomenal consciousness *as such*, that which is common (and peculiar) to all phenomenally conscious states. A comprehensive theory of phenomenal consciousness does require an account of qualitative character, because we also want to know what makes something the phenomenally conscious state it is given that it is one at all. But that is a secondary concern. The primary explanandum is what makes something a phenomenally conscious state at all—subjective character. Thus although Chap. 3 of SC is dedicated to an account of qualitative character, the bulk of the book is devoted to subjective character. Chap. 4 makes the case for a self-representational account of subjective character, and subsequent chapters refine and expand on that account by addressing specific dimensions thereof.

The fundamental argument for self-representationalism about subjective character can be set out in three stages. First, for a conscious experience to be not only *in* me, but also *for* me, I would have to be *aware* of it. The awareness in question is certainly quite special and elusive, and much should be said about what it is exactly. But *some* kind of awareness is necessary if the state is to exhibit for-me-ness. Moreover, once we fix on the right sort of awareness, it is also likely to be sufficient: in virtue of my special awareness of my experience, the experience is *for me* in the relevant sense. So we can reason as follows: a mental state of mine is phenomenally conscious iff it has for-me-ness (subjective character); a mental state has for-me-ness (subjective character) iff I am aware of it in the right way; therefore, a mental state of mine is phenomenally conscious iff I am aware of it in the right way. This is the first stage of the argument.

The second stage employs crucially a pair of relatively uncontroversial lemmas, to the effects that (a) being aware of something is a matter of representing it and (b) representing something is a matter of being in mental state that represents it. With these in place, we can argue that a phenomenally conscious state is represented by some mental state. The reasoning is this: I am aware of a mental state iff I represent it; I represent a mental state iff I am in a mental state that represents it; therefore, I am aware of a mental state of mine in the right way iff I am in a mental state that represents it in the right way. Taking into account the first stage of the argument, this leads to the thesis that a mental state of mine is phenomenally conscious iff I am in a mental state that represents it in the right way.

The third stage takes us from representation to self-representation. If a conscious state is always a represented state, the following question arises: are the conscious state and its representation numerically distinct or numerically identical? The third stage of the argument proceeds by offering considerations in favor of the latter option. More precisely: either it is always the case that a mental state of mine is phenomenally conscious iff it represents itself in the right way, or it is at least sometimes the case that a mental state of mine is phenomenally conscious if I am in a numerically distinct state that represents it; for a variety of reasons, it is never the case mental state is phenomenally conscious if I am in a numerically distinct state that represents it; therefore, it is always the case that a mental state is phenomenally

conscious iff it represents itself in the right way. This is, in effect, self-representationalism. The question, at this point, is what justifies the second premise in this last stage of the argument—what justifies the idea that conscious states are never conscious in virtue of being represented by numerically distinct states. In Chap. 4 of SC, I offer four arguments for this, but space limitations do not permit me to recite them here.

Putting together all these steps, we obtain a sustained argument for the thesis that a mental state is phenomenally conscious iff it represents itself in the right way. At this point it remains only to get clear on what the 'right way' is. This is something I am not very confident about, but what I propose in the book is that the relevant self-representation is non-derivative, specific, and essential. The first qualification is designed to rule out merely derivative, or conventional, self-representation (as in, e.g., the sentence 'this very sentence is written in black'); the second to rule out generic self-representation (as in, e.g., having the thought that all thoughts are neurally realized); the third to rule out purely accidental self-representation (as when I think that my mother's nieceless brother's only nephew is probably short before realizing that I am my mother's nieceless brother's only nephew).

An immediate challenge facing self-representationalism is to make sense of the notion of self-representation. This has two central parts: explicating what is involved in self-representation, and explaining how it can occur in the natural world. In Chap. 6 of SC, I offer an account of self-representation intended to make it consistent with naturalistic accounts of mental representation, which turns out to be a non-trivial exercise.

Once this general account is developed, we can seek brain structures and processes that implement it. Although an endeavor of this sort is extremely speculative at present, I indulge in it in Chap. 7 of SC. With the aid of several empirical claims, the speculative hypothesis I arrive at is this: phenomenally conscious states are realized by brain activations neurally synchronized with activation in the dorsolateral prefrontal cortex (dlPFC). Thus there is an intimate connection between the property of being phenomenally conscious and the property of being neurally synchronized with dlPFC activation.

The intimate connection can be construed as the correlation relation, or more strongly the realization relation, or more strongly yet the reduction relation. In Chap. 8, I try to remove familiar obstacles to the reduction of consciousness, thereby making the case that the property of being neurally synchronized with dlPFC activation is not only the neural *correlate* and *realizer*, but also a neural *reducer*, of consciousness. In particular, I argue that whether or not phenomenal consciousness actually *is* functionalizable, its self-representational nature entails that it is bound to *appear* uniquely resistant to functionalization. This is important, because functionalization is plausibly a necessary step to reduction, so a property that appeared uniquely resistant to functionalization would perforce appear uniquely resistant to reduction—again, regardless of whether it actually is. (See Chap. 8 of SC for details.) That is the final chapter of SC.