Journal of Cognition and Neuroethics

Anesthesia and Consciousness

Rocco J. Gennaro
University of Southern Indiana

Biography
Dr. Rocco J. Gennaro is the Philosophy Department Chairperson and a Professor of Philosophy at the University of Southern Indiana. He received his Ph.D. in 1991 at Syracuse University. Dr. Gennaro's primary research and teaching interests are in Philosophy of Mind/Cognitive Science (especially consciousness), Metaphysics, Early Modern History of Philosophy, NeuroEthics, and Applied Ethics. He has published ten books (as either sole author or editor) and over fifty articles and book chapters in these areas.

Publication Details
Journal of Cognition and Neuroethics (ISSN: 2166-5087). February, 2018. Volume 5, Issue 1.

Citation
Gennaro, Rocco J. 2018. "Anesthesia and Consciousness." *Journal of Cognition and Neuroethics* 5 (1): 49–69.

Anesthesia and Consciousness

Rocco J. Gennaro

Abstract

For patients under anesthesia, it is extremely important to be able to ascertain from a scientific, third-person point of view to what extent consciousness is correlated with specific areas of brain activity. Errors in accurately determining when a patient is having conscious states, such as conscious perceptions or pains, can have catastrophic results. Here, I argue that the effects of (at least some kinds of) anesthesia lend support to the notion that neither basic sensory areas nor the prefrontal cortex (PFC) is sufficient to produce conscious states. I also argue that it this is consistent with and supportive of the higher-order thought (HOT) theory of consciousness. I therefore disagree in some ways with Mehta and Mashour (2013), who argue that evidence from anesthesia mainly favors a first-order representational (FOR) theory, as opposed to HOT theory (and many other theories, for that matter).

Keywords

Anesthesia, Consciousness, Higher-Order Thought, Neural Correlates of Consciousness, Prefrontal Cortex

For patients under anesthesia, it is extremely important to be able to ascertain from a scientific, third-person point of view to what extent consciousness is correlated with specific areas of brain activity (Mashour 2010, Hutt and Hudetz 2015a). Errors in accurately determining when a patient is having conscious states, such as conscious perceptions or pains, can have catastrophic results. Here, I argue that the effects of (at least some kinds of) anesthesia lend support to the notion that neither basic sensory areas nor the prefrontal cortex (PFC) is sufficient to produce conscious states. I also argue that it this is consistent with and supportive of the higher-order thought (HOT) theory of consciousness (Rosenthal 2005, Gennaro 2012). I therefore disagree in some ways with Mehta and Mashour (2013) who argue that evidence from anesthesia mainly favors a first-order representational (FOR) theory as opposed to HOT theory (and many other theories, for that matter).

1. Introduction and Terminology

Perhaps the most fundamental and commonly used notion of "conscious" is captured by Thomas Nagel's famous "what it is like" sense (Nagel 1974). When we are in a conscious mental state, there is "something it is like" for us to be in that state from the subjective or first-person point of view. When we smell a rose or have a conscious visual experience, there is something it "seems" or "feels like" from our own perspectives. An organism such as a bat is conscious if it is able to experience the world through its echolocation senses. There is also something it is like to be a conscious creature, whereas there is nothing it is like to be a table or tree. This is primarily the sense of "conscious state" used throughout this article. "What it's like" basically means "how a conscious state is for the subject."

Let us also keep in mind the distinction between *state* and *creature* consciousness (Rosenthal 1993). We sometimes speak of an individual mental state, such as a pain or perception, as being conscious. On the other hand, we also often talk about organisms or creatures as conscious, such as when we say that "human beings are conscious" or "cats are conscious." Creature consciousness is simply meant to refer to the fact that an organism is awake, as opposed to sleeping or in a coma. However, some kind of state consciousness is normally implied by creature consciousness; that is, if a creature is conscious, then it must have conscious mental states. Perhaps there are cases where one is state conscious but not creature conscious, such as when one is having a vivid dream. Another possible case is "locked-in syndrome," which is a medical condition where brain damage has affected only motor functions and leaves the patient immobile and unresponsive to stimuli but consciousness remains normal. Mashour and LaRock (2008) refer to locked-in syndrome as the "inverse zombie problem," that is, cases of internally experienced consciousness without any behavioral sign, as opposed to the hypothetical philosopher's "zombie," who is not conscious but behaves in a manner indistinguishable from a conscious human. In this and other related troubling cases, such as persistent vegetative states or minimally conscious states, significant ethical concerns also frequently arise (Braddock 2017).

2. Some Evidence from Anesthesia

The overall available evidence strongly suggests that anesthesia (such as propofol and ketamine) primarily causes the suppression of "feedback" and "top-down" brain mechanisms or connectivity (e.g., Hudetz 2012; Lee et. al. 2013; Crone 2017). It will be helpful to review in some detail the findings and conclusions recently reached by experts in the field. For example, Hudetz explains that "networks based on the posterior parietal-cingulate-precuneus region as a hub and on the nonspecific thalamus are putative candidates for the neural correlate of the state of consciousness [NCCs]" (Hudetz 2012, 299). These areas, he notes, are "prime candidates for the functional networks of the forebrain that play a critical role in maintaining the state of consciousness" (Hudetz 2012, 291). Notice that there is no specific mention of the PFC or basic sensory areas as loci for the NCC.

In a similar fashion, Schrouff et al. (2011, 203) tell us that "results show that deep sedation [due to propofol] was associated with reduced interactions between all… associative cortices…However…the functional interactions of parietal areas were deteriorated to a significantly larger extent than those of frontal or temporal areas." Alkire, Hudetz, and Tononi (2008, 876) explain that "unconsciousness is likely to ensue when a complex of brain regions in the posterior parietal area is inactivated." Thus, there is continued emphasis on parietal brain areas as central to when a subject is having a conscious experience.

Crone et al. (2017) likewise do not focus on the PFC but rather conclude that:

> the data show that loss of consciousness, at least in the context of propofol-induced sedation, is marked by a breakdown of corticopetal projections from the globus pallidus. Effective connectivity between the globus pallidus and the ventral posterior cingulate cortex…fades in the transition from lightly sedated to full loss of consciousness and returns gradually as consciousness recovers. (Crone et al. 2017, 2727)

Further, Hutt and Hudetz (2015b) summarize Blain-Moraes et al. (2014) by explaining that in surgical patients "anesthetic-invariant electroencephalographic effects occur in cortical top-down connectivity. Specifically, ketamine is found to suppress fronto-parietal functional and directional connectivity, similar to that produced by propofol" (Hutt and Hudetz 2015b, 4). They conclude that

> the formerly favored bottom-up mechanisms of anesthetic action focusing on subcortical arousal centers and ascending thalamocortical information transfer are contrasted with the more recent cortical top-down explanations that are inherent to conscious perception and appear to be the preferential target of anesthetic modulation. Substantial electrophysiological and neuroimaging evidence from animal and human investigations supports the top-down mechanisms as a causally sufficient explanation for anesthetic-induced unconsciousness. (Hutt and Hudetz 2015b, 4)

Hudetz and Mashour (2016) frame the matter in the following way:

> After a … dose of propofol, highly connected "hubs" in the brain undergo a reconfiguration, with connectivity patterns in the posterior

> parietal cortex being disrupted. Much like an airport system, a disrupted hub would entail a reduction in incoming traffic, which is exactly what is observed in the form of reduced communication from frontal cortex to posterior parietal cortex. (Hudetz and Mashour 2016, 1233)

Many of the above authors also emphasize the need for necessary connectivity and interaction between different areas of the brain. In addition, Boveroux et al. (2010) show that anesthetic-induced unconsciousness is *not* correlated with inactivation of primary sensory cortical areas. Transverse and sagittal sections of primary visual and the auditory cortices during wakefulness and propofol-induced unconsciousness show the relative preservation across states. That is, neural activity in the primary visual and auditory cortices is preserved while the patient is under anesthesia. Thus, it seems that whatever generally makes a visual or auditory mental state conscious cannot be within the primary cortices. Much the same seems to be the case for various other kinds of conscious states, such as emotions and pains. Indeed, fear and pain, for example, seem to essentially involve an emotional element or cognitive attitude (Baars and Gage 2010, Ch. 13). The limbic system, for example, contains some subcortical structures (such as the amygdala and hypothalamus) as well as some cortical structures (such as the cingulate gyrus). Indeed, the neural realization of pain and emotion is fairly complex and also distributed in different brain areas. In any case, the overall idea is that for normal conscious states, there are "lower" areas of brain activity accompanied higher-level top down cortical interaction. As we have seen, anesthesia mainly disrupts the top-down neural activity.

3. Support for HOT Theory

I think that the above, in turn, provides some support for the view that having conscious states requires having *higher-order thoughts* (HOTs) most often located in-between early sensory areas and the prefrontal cortex (PFC). The HOT theory of consciousness says that what makes a mental state conscious is that there is a suitable higher-order thought directed at the mental state (Rosenthal 2005; Gennaro 2012). HOTs are "meta-psychological" or "metacognitive" states, that is, mental states directed at other mental states. HOT theory is primarily concerned with explaining how conscious mental states differ from unconscious mental states.

Let's back up for a moment. A central question which should be answered by any theory of consciousness is: What makes a mental state a *conscious* mental state? That is, how do we distinguish between unconscious mental states and conscious mental states? HOT theorists put significant initial weight on what has come to be known as the transitivity principle (TP).

> TP: Conscious states are mental states that I am "aware of" in some sense.

TP seems intuitively true and perhaps even true by definition. When I am in a conscious visual state, I am aware of being in that state. On the other hand, *unconscious* states are those mental states of which I am not aware. I am not aware of being in my current unconscious states. If I am having a subliminal perception, then I am not aware of being in that state. For various reasons, many (including myself) hold that such "meta-awareness" is best understood as a thought composed of concepts, as opposed to, say, a perception.[1]

There is an important and very relevant additional subtlety to HOT theory, however. When a conscious mental state is a first-order world-directed state the higher-order thought (HOT) is *not* itself conscious; otherwise, circularity and an infinite regress would follow. In such cases, we are unaware of the HOTs themselves since our conscious focus in world-directed. But when the HOT is itself conscious, there is a yet higher-order (or third-order) thought directed at the second-order state. In this case, we have *introspection* which involves having a conscious HOT directed at an inner mental state. When one introspects, one's attention is directed back into one's mind, but when one has a first-order conscious state one's attention is outer-focused (see figure 1).

My view is that HOTs, especially the *un*conscious HOTs that accompany first-order conscious states, need not occur in the prefrontal cortex (PFC) area. This also seems to be supported by recent work on anesthesia, as we have already seen to some extent (more on this below). So, although HOT theory demands that conscious states be distributed to some degree in the brain (i.e., beyond basic sensory areas), I opt for a more moderate view with more limited neural connections required (e.g., recurrent feedback loops), especially with respect to first-order conscious states.[2]

1. HOT theory is most often contrasted with HOP (higher-order perception) theory. See e.g. Gennaro (2004), Rosenthal (2004), and Gennaro (2012, chapter 3) for much more discussion of alternative HO theories.

2. Actually, I prefer to treat the lower-order state and the unconscious HOT as parts of a single complex unified state. This is a position I have called the "wide intrinsicality view" or WIV (Gennaro, 1996, 2006, 2012).

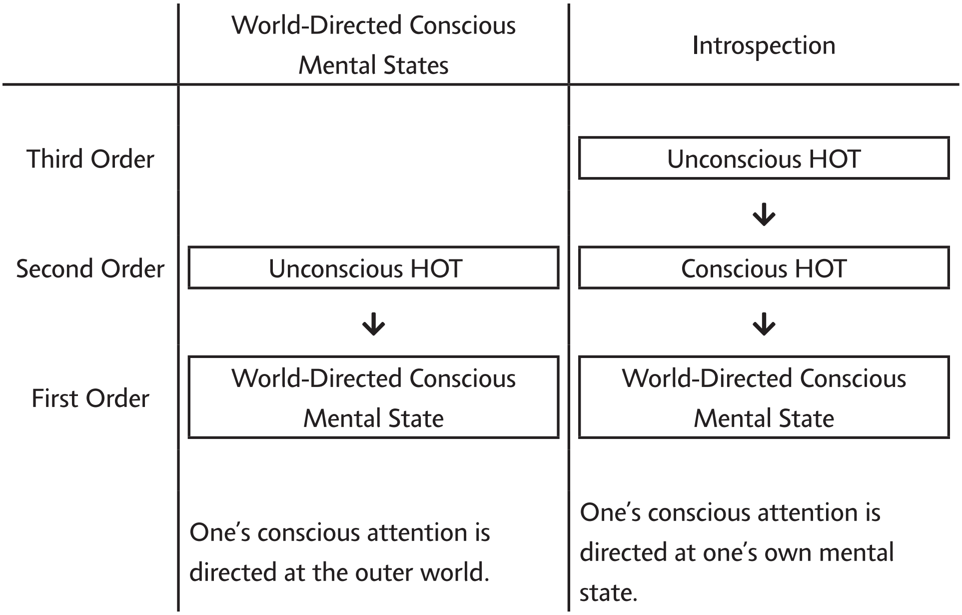

Figure 1. The Higher-Order Thought (HOT) Theory of Consciousness

At the neural level, then, we might borrow from Gerald Edelman and others who have argued that *feedback loops* (or "reentrant pathways" or "back projections") in the neural circuitry of the brain are essential for conscious awareness (Edelman and Tononi 2000a; 2000b). As Patricia Churchland once put it, "The idea is that some neurons carry signals from more peripheral to more central regions…while others convey more highly processed signals in the reverse direction. . .It is a general rule of cortical organization that forward-projecting neurons are matched by an equal or greater number of back-projecting neurons" (2002, 148–149). The brain structures involved in loops seem to resemble the structure of at least some form of HOT theory; namely, lower-order and higher-order states mutually interacting to produce conscious states. Edelman and Tononi, for example, sometimes emphasize the global nature of conscious states, and it is reasonable to interpret this as the view that conscious states comprise both the higher-and lower-order states. What they call the "dynamic core" is generally "spatially distributed and thus cannot be localized to a single place in the brain" (Edelman and Tononi 2000a, 146).

Other related support comes from Victor Lamme (2003; 2004) who argues that recurrent processing is necessary before the properties of an object are attentively grouped and the stimulus can enter consciousness. Based on experimental results, such as texture segregation and visual search tasks, Lamme argues that the so-called "feedforward sweep" is not sufficient for consciousness. Lamme also explains that "backward masking" renders a visual stimulus invisible by presenting a second stimulus shortly after the first (about 40 milliseconds later, but perhaps up to 110 msecs). Nonetheless, the masked (invisible) stimulus still evokes significant feedforward activation in visual and even nonvisual areas. It seems that the feedback interaction from higher to lower visual areas is suppressed by backward masking, thereby disrupting reentrant processing and inhibiting the production of a conscious states (Fahrenfort, Scholte, and Lamme 2007; Kouider and Dehaene 2007). This suggests that neural activity beyond basic sensory areas is necessary in order to have a conscious perceptual state.

To use one nonvisual example, consider tactile awareness in the somatosensory cortex, extensively reviewed in Gallace and Spence (2010). Once again, there seems to be evidence of feedback activity from higher brain areas necessary for conscious tactile experiences. Gallace and Spence explain that "activation of early sensory areas is insufficient to sustain awareness of tactile sensations. . . . Higher order structures seem necessary" (2010, 50). So, we might say, that tactile information becomes conscious when earlier somatosensory areas trigger a feedback signal from a higher-order representation.

4. Mehta and Mashour's Argument

I therefore disagree to some extent with Mehta and Mashour (2013; M&M hereafter), who argue that evidence from anesthesia mainly favors a first-order representational (FOR) theory, as opposed to HOT theory (and many other theories, for that matter). A FOR theory of consciousness is one that attempts to explain and reduce conscious experience primarily in terms of world-directed (or first-order) intentional states (e.g. Tye 2000, Dretske 1995).

M&M start out by noting that

> a complete theory of consciousness must explain…*first* …what we term *general consciousness*: what makes a state conscious *at all*, as opposed to wholly unconscious. This should explain the difference between, e.g., one's conscious state when one sees a red thing and one's unconscious state when one is anesthetized…. *Second*…*specific consciousness*: what gives a state its specific *phenomenal character* [or

> "content"], rather than some alternative phenomenal character. This… should explain the difference between one's conscious state when one sees something red and one's conscious state when one sees something green (or hears a loud noise, or feels pain). (Mehta and Mashour 2013, 2)

M&M also define two broad types of systems that participate in conscious processing, namely

> *sensory systems* [which] are dedicated to the detection of highly specific perceptible features [which] may be tuned to modality-specific properties (such as color and tone) or properties detectible via multiple modalities (such as motion and spatial location)…[and] *post-sensory systems* [which] perform a broad variety of functions, including modulation of sensory processing via top-down attention... (Mehta and Mashour 2013, 2)

But then, M&M tie HOTs mainly to the prefrontal cortex (PFC) when they say the following:

> Lau and Rosenthal (2011) hypothesize that higher-order representations are harbored in post-sensory systems, such as posterior parietal regions and *especially the dorsolateral prefrontal cortex*, while first-order representations are harbored in sensory systems. Although this neural interpretation is not forced on higher-order representationalists, we will henceforth adopt it because (a) several higher-order representationalists have endorsed this interpretation, and (b) higher-order representationalism is very difficult to test scientifically without some neural interpretation. Given this interpretation, higher-order representationalists identify post-sensory areas as the neural correlates of *general* consciousness; they identify post-sensory and perhaps also sensory areas as the neural correlates of *specific* consciousness (Mehta and Mashour 2013, 3, my emphases).

So, it turns out that one main reason that M&M favor FOR over HOT theory is that they suppose that all (most?) HOT theorists think that HOTs mainly occur in "post sensory systems…*especially the dorsolateral prefrontal cortex*." But, when we look closely at M&M's own paper, we can see that much of the evidence they cite, at minimum, equally favors HOT theory and the same surely goes for the other anesthesia evidence cited earlier. That is, anesthesia primarily targets post-sensory areas but not the PFC. One would expect HOTs to occur in post-sensory areas. To be fair, M&M do ultimately concede that a modified HOT theory could accommodate the data.

As M&M rightly acknowledge, and as we saw earlier, there is other supportive evidence from Lamme and colleagues. M&M ask us to

> consider visual processing, which includes both a fast *feedforward sweep* and slower *recurrent processing* loops. In the feedforward sweep, which is completed in about 100ms, activation proceeds in a swift, unidirectional cascade from the retina, to the lateral geniculate nucleus of the thalamus, to V1 in the occipital cortex, to higher visual processing areas (includingV2, V3, V4, and V5), and finally to more rostral structures. By contrast, recurrent processing loops involve reciprocal information transfer between the cortical areas, and corresponding thalamic regions via horizontal and feedback connections. Recurrent processing may occur at many scales, from a local scale (within a sensory modality) to a global scale (implicating executive function or spanning different sensory modalities). (Mehta and Mashour 2013, 2-3)

Still, M&M do rightly suppose that "…the simplest explanation of the current data is that…post-sensory processing regions *alone* are the neural correlates of general consciousness. [But then they say that] "this result favors first-order representationalism and higher-order representationalism" (Mehta and Mashour 2013, 6). It seems to me, however, that it favors the higher-order approach much more so than FOR. As a matter of fact, Mashour (2014) elsewhere himself says that "there is growing evidence… that general anesthetics disrupt higher-order cognitive processes and that networks of association cortex may be particularly susceptible to anesthetic effects…" (Mashour 2014, 7). Similarly, he explains that "consciousness and anesthetic-induced unconsciousness are associated with multimodal association cortex rather than primary sensory cortex" (Mashour 2014, 2).

It is somewhat unclear to me how exactly a FOR theorist can suppose that post-sensory areas constitute the locus of "general" consciousness. Presumably, it is because M&M understand FOR theories to include the notion that first-order conscious states must be "available to" post-sensory areas, that is, a conscious representation must be "available to" the subject as a reason for action and belief formation. They explain as follows:

> according to [FOR] theory, consciousness is hypothesized to consist in (i) first-order representations directed at the world which (ii) are directly available to the subject for action selection, belief formation, planning, etc. Condition (i) embodies an approach to specific consciousness: the specific phenomenal character of a representation is determined wholly by its content. Condition (ii) complements this with an approach to general consciousness: for a representation to be conscious rather than unconscious is for it to be directly available to the subject for action selection, belief formation, planning, etc. (Mehta and Mashour 2013, 6).

But it is difficult to see how the mere "availability" to the subject for various purposes can make a representation conscious, at least as opposed to an *actual* HOT directed at the first-order representation. Similarly, it is unclear just how such a disposition can confer *actual* consciousness on an otherwise unconscious mental state. Further, and more to the point, when an anesthetic suppresses post-sensory activity, it is so much clearer that HOTs cannot occur in those areas and thus conscious states will not occur. On the other hand, what does it mean to say that the "availability" to the subject has been suppressed? How does an unconscious first-order representation "know" that such availability has been cut off? Isn't the first-order state still "potentially" conscious in some sense? Why aren't sensory areas enough to produce conscious states if sensory area mental states are merely disposed to be made aware of by post-sensory states?[3]

It is worth briefly noting here that Flohr (2000) has argued that N-methyl-D-aspartate (NMDA)-mediated transient neural assemblies are essential to consciousness based on evidence from anesthesia. Flohr also cites HOT theory approvingly as a way to explain the overall structure of conscious states and the effects of anesthesia. The idea is that anesthetics destroy conscious mental activity because they interfere with the functioning of NMDA receptors. According to Flohr, the activation of the NMDA system is necessary for the mechanisms underlying consciousness. Flohr explicitly relates his theory to higher-order accounts of consciousness by arguing that the NMDA synapse implements the binding mechanism that the brain uses to produce widely distributed representations (Flohr 2000, 252–253). One potential problem with Flohr's account might be that he is focused too narrowly on overall creature consciousness in the sense of

3. Thanks to Benedicte Veillet for pressing this point. For more on these themes and different FOR and HOT theories, see Rosenthal 2004 and Gennaro 2012, chapter 3.

a creature being awake or aware of its surroundings (as opposed to state consciousness). In some ways, however, his emphasis makes sense when thought of from the point of view of anesthesia and neurochemistry. The question is indeed often about whether or not the *patient* is unconscious or when the *person* loses consciousness. Of course, we still want to know if the patient is experiencing any conscious states, and especially pains.

5. HOT Theory and the PFC

If we are correct thus far and anesthesia does not mainly target PFC areas, it would seem that a HOT theorist should look elsewhere for where at least many HOTs occur in the brain. Thus, I have argued elsewhere (Gennaro 2012, ch. 9) that HOT theory need not be committed to the view that the PFC is required for having conscious states, contrary to Kriegel (2007) and Block (2007) who (like Lau and Rosenthal) also suppose that HOT theorists hold that HOTs are realized in the PFC. Still, it is likely true that the PFC is required for the more sophisticated *introspective* states, but this is not a problem for HOT theory because it does not require introspection for merely having first-order (outer-directed) conscious states (see figure 1 again).

On Kriegel's theory, for example, three "elements" are required for NCCs:

1. A "floor-level" (or first-order) representation,
2. A "higher-order" representation of (1), and
3. The "functional integration" of (1) and (2) into a single unified state via some binding mechanism.

The likely NCCs for the floor-level representations will depend on the modality, such as V1-V5/MT for perceiving a moving patch of blue color. According to his view, this has to do with the *contents* of consciousness (or "specific" consciousness, using M&M's terminology), as opposed to consciousness *as such* (or "general" consciousness). So far so good, but then Kriegel says that the likely NCCs for the second-level or higher-order representations are in the PFC. In reply, we should point out that Kriegel's discussion of his "second element" reflects some very sophisticated abilities, such as "executive functions" and "attentional control," which are better understood as *introspective* capacities. Again, it might very well be that *these* higher cognitive capacities are indeed subserved by PFC activity but there is no reason to think that they are required for having first-order conscious mental states (even according to HOT theory).

Block (2007, 485) also states that "since frontal areas are likely to govern higher-order thought, low frontal activity in newborns [and most animals?] may well indicate a lack of higher order thoughts about genuine sensory experiences." Although I agree with Block that PFC activity is not necessary for having first-order conscious states, I disagree with the claim that "frontal areas are likely to govern higher-order thought," unless he primarily means *introspection*; that is, conscious HOTs. In short, a HOT theorist is not be committed to the view that PFC activity is required for having all conscious states. This is also important in order to counter the frequently made charge that HOT theory rules out infant and (most) animal consciousness (Seager 2004). Indeed, one HOT theorist has accepted this otherwise undesirable consequence of HOT theory (Carruthers 2000, 2005).[4] It is also clear from the evidence adduced earlier that neural activity in post-sensory areas (but not including the PFC) result in first-order conscious states and are responsible for general consciousness. After all, it is suppression of activity in those areas which eliminate consciousness.

In addition to the anesthesia-based evidence cited thus far, there are independent reasons to think that conscious states occur without the PFC. For example, conscious experience is not eliminated entirely when there is extensive PFC damage, even in lobotomies (Pollen 2003). And when subjects are engaged in a perceptual task or absorbed in watching a movie, there is widespread neural activation but little PFC activity (Goldberg, Harel, and Malach 2006). Although other studies do show *some* PFC activation in similar experiments, this is mainly because of the need for subjects to *report* their experiences and the PFC is likely to be activated when there is *reflection* or *introspection* about one's experiences.

But is there any positive reason to think that unconscious HOTs can occur outside of the PFC? I think there is. Assuming that HOTs can be understood as a form of self-consciousness, as seems reasonable I think, unconscious HOTs might then be regarded as a kind of "pre-reflective" self-consciousness (as opposed to reflective or introspective self-consciousness). Newen and Vogeley (2003), for example, go so far as to distinguish five levels of self-consciousness ranging from "phenomenal self-acquaintance" and "conceptual self-consciousness" up to "iterative meta-representational self-consciousness." Citing numerous experiments, they point to various "neural signatures" of self-consciousness, but the PFC is rarely mentioned, and then, usually only with regard to the more sophisticated forms of self-consciousness. Other brain areas are much more prominently

4. But see e.g. Gennaro 2009, Gennaro 2012, chapters 7 and 8, for my most recent attempts to rebut that line of argument.

identified, such as the medial and inferior parietal cortices, the temporo-parietal cortex, the posterior cingulate cortex, and the anterior cingulate cortex.

Even when considering the neural signatures of "theory of mind" and "mindreading," Newen and Vogeley cite experiments indicating that such meta-representation is best located in the anterior cingulate cortex and also showed activation in the right temporo-parietal junction and the medial aspects of the superior parietal lobe. Related, and more recent, support for this position can be found, for example, in the work of Rebecca Saxe, who has extensively studied brain regions most associated with thinking about other people's thoughts, sometimes called "mindreading" and involving a so-called "theory of mind" (Saxe 2009; 2010). The temporo-parietal junction, the posterior cingulate, the medial precuneus, and parts of the temporal sulcus are identified as the primary sites for this kind of cognition.

6. Worries about NCCs?

One might wonder if we are playing too fast and loose with talk of NCCs. At the least, it is important to avoid several problems and potential pitfalls when discussing evidence related to NCCs. For example, one issue is determining exactly how the NCC is related to consciousness. Although a case can be made that many NCCs are *necessary* for conscious mentality, it is sometimes unclear if they are *sufficient*. For one thing, many candidates for NCCs can also occur unconsciously, such as feedback loops in earlier sensory areas. Second, there are obviously other background conditions that must obtain (e.g., breathing, proper blood flow) in order for a given NCC to suffice for consciousness. Even pinning down a narrow-enough necessary condition is not as easy as it might seem.

A related worry has to do with the very use of the term "correlate." As any philosopher, scientist, and even undergraduate student should know, saying that "A is correlated with B" is rather weak by itself (though it can be an important first step), especially if one wishes to establish a stronger *causal* or *identity* claim between consciousness and neural activity. Even if a solid correlation can be established, we cannot automatically conclude that there is an identity relation. One might even suppose that the search for NCCs is somewhat neutral with respect to the metaphysics of mind, though a materialist might urge us at some point to accept an identity claim on the basis of the principle of simplicity. Still, perhaps A causes B or B causes A, and that's why we find the correlation. Maybe there is even some *other* neural process C that causes both A and B.

Chalmers (2000) presents several useful distinctions and definitions for the purpose of conceptual clarity (cf. Block 2007; Hohwy 2007). For one thing, we should distinguish between having the conscious mental state itself (or "vehicle") and its content. Thus, Chalmers presents the following definitions:

> A *content* NCC is a neural representational system N such that the content of N directly correlates with the content of consciousness. (Chalmers 2000, 20; italics mine).
>
> A *state* N1 of system B is a neural correlate of phenomenal property P if N's being in N1 directly correlates with the *subject* having P. (22; italics mine).

In our discussion of anesthesia, we must be careful not to suppose that post-sensory brain areas determine what M&M call "specific" consciousness; that is, the *content* NCCs of conscious perceptual states. Actually, some HOT theorists think that it takes *both* sensory and post-sensory areas to produce *specific* consciousness in the sense that the first-order content must be properly referenced or matched by the HOT's content (Gennaro 2012).[5] In any case, we have already seen that general consciousness is not realized in basic sensory areas and is likely to be in various post-sensory areas. This is more like what Chalmers has in mind by "state" NCC as opposed to the "content" NCC.

It is then important to recognize that any interesting NCC would at least need to isolate the *minimal* brain area responsible for a conscious state. Thus, one finds the following:

> An NCC is a *minimal neural system* N such that there is a mapping from states of N to states of consciousness, where a given state of N is sufficient, under conditions C, for the corresponding state of consciousness. (Chalmers 2000, 31; italics mine).

From the evidence we have examined regarding anesthesia, it certainly seems as if the minimal neural system N in question would include both sensory and post-sensory areas when one is having a conscious first-order perceptual state. Each area is necessary for conscious states, but they are jointly sufficient for the relevant state of consciousness. To be more specific with regard to post-sensory areas, the posterior parietal and cingulate cortices seem to be especially important.

5. There is significant disagreement on this matter among HOT theorists, such as what would happened if or when a HOT's content misrepresents its target mental state or even when there is no target state at all. See Gennaro (2012, chapters 4 and 6) for some discussion.

Others make similar remarks and distinctions with respect to NCCs, such as when Block (2007, 489) explains that a "minimal neural basis is a necessary part of a neural sufficient condition for conscious experience," and when Koch (2004, 16) tells us that the NCC is "the minimal set of neuronal events and mechanisms jointly sufficient for a specific conscious percept." The main point is to find a neural correlation that is a reasonably interesting subset of the entire brain activity at a given time. It would be much less informative, and perhaps just trivial, to learn that the *entire* brain is sufficient for having a conscious state. In a similar vein, one might distinguish between the *core* and *total* NCC. The *core* neural basis of a conscious state is the part of the total neural basis that distinguishes conscious states from states with other conscious contents. This is very much like what M&M call "specific" consciousness and what Chalmers calls the "content NCC." The *total* neural basis of a conscious state is itself sufficient for the instantiation of that conscious state (Block 2007, 482). Once again, this seems to involve both sensory and post-sensory areas. We also need to distinguish the NCC from what might be viewed as other "enabling conditions," which refer to other aspects of a functioning body, such as proper blood flow and functioning lungs and heart (see Block 2007, 485-486, for some discussion). There may be deeper problems here as well and perhaps we even need new experimental approaches (Hohwy 2009). At the least, it is also crucial to design experiments with controls such that the *only* difference between a pair of trials is the presence of consciousness. We can then use, say, functional magnetic resonance imaging (fMRI) to ascertain any neural differences between such cases. Neurophysiologists sometimes use the "subtraction method;" namely, subtract control or constant background neural activity from the neural activity of a given task which involves conscious perception.

7. Conclusion

The effects of (at least some kinds of) anesthesia lend support to the notion that neither basic sensory areas nor the prefrontal cortex (PFC) is sufficient to produce conscious states. Rather, it takes the combination of sensory areas and post-sensory areas (not including the PFC) in order for there to be a first-order conscious state. It was also argued that it this is consistent with and most supportive of the higher-order thought (HOT) theory of consciousness. Still, we must be careful as to how to characterize the NCCs in question.

References

Alkire, Michael, Anthony Hudetz, and Giulio Tononi. 2008. "Consciousness and Anesthesia." *Science* 322 (5903): 876–880.

Baars, Bernard, and Nicole Gage. 2010. *Cognition, Brain, and Consciousness, 2nd edition*. San Diego: Elsevier/Academic Press.

Blain-Moraes, Stefanie, UnCheol Lee, SeungWoo Ku, GyuJeong Noh, and George Mashour. 2014. "Electroencephalographic Effects of Ketamine on Power, Cross-frequency Coupling, and Connectivity in the Alpha Bandwidth." *Frontiers in Systems Neuroscience* 8 (July): 1-9.

Block, Ned. 2007. "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience." *Behavioral and Brain Sciences* 30 (5): 481-499.

Boveroux, Pierre, Audrey Vanhaudenhuyse, Marie-Aurelie Bruno, Quentin Noirhomme, S. Lauwick, and A. Luxen, Christian Degueldre, Alain Plenevaux, Caroline Schnakers, Christophe Phillips, Jean-Francois Brichant, Vincent Bonhomme, Pierre Maquet, Michael D. Greicius, Steven Laureys, and Melanie Boly. 2010. "Breakdown of Within- and Between-Network Resting State Functional Magnetic Resonance Imaging Connectivity during Propofol-Induced Loss of Consciousness." *Anesthesiology* 113 (5): 1038-1053.

Braddock, Matthew. 2017. "Should We Treat Vegetative and Minimally Conscious Patients as Persons?" *Neuroethics* 10 (2): 267-280.

Carruthers, Peter. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge University Press.

Carruthers, Peter. 2005. *Consciousness: Essays from a Higher-Order Perspective*. New York: Oxford University Press.

Chalmers, David. 2000. "What is a Neural Correlate of Consciousness?" In *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, edited by Thomas Metzinger, 17-39. Cambridge: The MIT Press.

Churchland, Patricia. 2002. *Brain-Wise*. Cambridge: The MIT Press.

Crone, Julia, Evan Lutkenhoff, Branden Bio Steven Laureys, and Martin Monti. 2017. "Testing Proposed Neuronal Models of Effective Connectivity within the Cortico-basal Ganglia-thalamo-cortical Loop During Loss of Consciousness." *Cerebral Cortex* 27 (4): 2727-2738.

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

Edelman, Gerald, and Giulio Tononi. 2000a. "Reentry and the Dynamic Core: Neural Correlates of Conscious Experience." In *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, edited by Thomas Metzinger, 139-151. Cambridge: The MIT Press.

Edelman, Gerald, and Giulio Tononi. 2000b. *A Universe of Consciousness*. New York: Basic Books.

Fahrenfort, Johannes, Steven Scholte, and Victor Lamme. 2007. "Masking Disrupts Reentrant Processing in Human Visual Cortex." *Journal of Cognitive Neuroscience* 19 (9): 1488-1497.

Flohr, Hans. 2000. "NMDA Receptor-Mediated Computational Processes and Phenomenal Consciousness." In *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, edited by Thomas Metzinger, 245-258. Cambridge: The MIT Press.

Gallace, Alberto, and Charles Spence. 2010. "Touch and the Body: The Role of the Somatosensory Cortex in Tactile Awareness." *Psyche* 16 (1): 31-67.

Gennaro, Rocco. 1996. *Consciousness and Self-Consciousness: A Defense of the Higher-Order Thought Theory of Consciousness*. Amsterdam: John Benjamins.

Gennaro, Rocco. 2004. *Higher-Order Theories of Consciousness: An Anthology*. Amsterdam: John Benjamins.

Gennaro, Rocco. 2006. "Between Pure Self-Referentialism and the (Extrinsic) HOT Theory of Consciousness." In *Self-Representational Approaches to Consciousness*, edited by Uriah Kriegel and Ken Williford, 221-248. Cambridge: The MIT Press.

Gennaro, Rocco. 2009. "Animals, Consciousness, and I-thoughts." In *Philosophy of Animal Minds*, edited by Robert Lurz, 184-200. New York: Cambridge University Press.

Gennaro, Rocco. 2012. *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts*. Cambridge: The MIT Press.

Goldberg, Ilan, Michael Harel, and Rafael Malach. 2006. "When the Brain Loses Its Self: Prefrontal Inactivation during Sensorimotor Processing." *Neuron* 50 (2): 329-339.

Guldenmund, Pieter, Ithabi S. Gantner, Katherine Baquero, Tushar Das, Athena Demertzi, Pierre Boveroux, Vincent Bonhomme, Audrey Vanhaudenhuyse, Marie-Aurélie Bruno, Olivia Gosseries, Quentin Noirhomme, Muriëlle Kirsch, Mélanie Boly, Adrian M. Owen, Steven Laureys, Francisco Gómez, and Andrea Soddu. 2016. "Propofol-Induced Frontal Cortex Disconnection: A Study of Resting-State Networks, Total Brain Connectivity, and Mean BOLD Signal Oscillation Frequencies." *Brain Connectivity* 6 (3): 225-237.

Hohwy, Jakob. 2007. "The Search for Neural Correlates of Consciousness." *Philosophy Compass* 2 (3): 461-474.

Hohwy, Jakob. 2009. "The Neural Correlates of Consciousness: New Experimental Approaches Needed?" *Consciousness and Cognition* 18 (2): 428-438.

Hudetz, Anthony. 2012. "General Anesthesia and Human Brian Connectivity." *Brain Connectivity* 2 (6): 291-302.

Hudetz, Anthony and George Mashour. 2016. "Disconnecting Consciousness: Is There a Common Anesthetic End Point?" *Anesthesia and Anelgesia* 123 (5): 1228-1240.

Hutt, Axel and Anthony Hudetz. eds. 2015a. *General Anesthesia: From Theory to Experiments*. Lausanne: Frontiers Media (*Frontiers in Systems Neuroscience*).

Hutt, Axel and Anthony Hudetz. 2015b. "Editorial: General Anesthesia: From Theory to Experiments." In *General Anesthesia: From Theory to Experiments*, edited by Axel Hutt and Anthony Hudetz, 4-6. Lausanne: Frontiers Media (*Frontiers in Systems Neuroscience*).

Koch, Christof. 2004. *The Quest for Consciousness: A Neurobiological Approach*. Englewood, CO: Roberts.

Kouider, Sid, and Stanislas Dehaene. 2007. "Levels of Processing During Non-conscious Perception: A Critical Review of Visual Masking." *Philosophical Transactions of the Royal Society of London, B* 362 (1481): 857-875.

Kriegel, Uriah. 2007. "A Cross-Order Integration Hypothesis for the Neural Correlate of Consciousness." *Consciousness and Cognition* 16 (4): 897-912.

Lamme, Victor. 2003. "Why Visual Attention and Awareness Are Different." *Trends in Cognitive Sciences* 7 (1): 12-18.

Lamme, Victor. 2004. "Separate Neural Definitions of Visual Consciousness and Visual Attention: A Case for Phenomenal Awareness." *Neural Networks* 17 (5): 861-872.

Lau, Hakwan, and David Rosenthal. 2011. "Empirical Support for Higher-Order Theories of Conscious Awareness." *Trends in Cognitive Sciences* 15 (8): 365-373.

Lee, UnCheol, SeungWoo Ku, GyuJeong Noh, SeungHye Baek, ByungMoon Choi, and George Mashour. 2013. "Disruption of Frontal-Parietal Communication by Ketamine, Propofol, and Sevoflurane." *Perioperative Medicine* 118 (6): 1264-1276.

Mashour, George. ed. 2010. *Consciousness, Awareness, and Anesthesia*. Cambridge: Cambridge University Press.

Mashour, George. 2014. "Top-down Mechanisms of Anesthetic-Induced Unconsciousness." *Frontiers of Systems Neuroscience* 8 (June): 1-10.

Mashour, George, and Eric LaRock. 2008. "Inverse Zombies, Anesthesia Awareness, and the Hard Problem of Unconsciousness." *Consciousness and Cognition* 17 (4): 1163-1168.

Mehta, Neil and George Mashour. 2013. "General and Specific Consciousness: A First-Order Representationalist Approach." *Frontiers in Psychology* 4 (July): 1-9.

Nagel, Thomas. 1974. "What is it Like to be a Bat?" *Philosophical Review* 83 (4): 435-456.

Newen, Albert, and Kai Vogeley. 2003. "Self-Representation: Searching for a Neural Signature of Self-Consciousness." *Consciousness and Cognition* 12 (4): 529-543.

Pollen, Daniel. 2003. "Explicit Neural Representations, Recursive Neural Networks and Conscious Visual Perception." *Cerebral Cortex* 13 (8): 807-814.

Rosenthal, David. 1993. "State Consciousness and Transitive Consciousness." *Consciousness and Cognition* 2 (3): 355-363.

Rosenthal, David. 2004. "Varieties of Higher-Order Theory." In *Higher-Order Theories of Consciousness: An Anthology*, edited by Rocco Gennaro, 17-44. Amsterdam: John Benjamins.

Rosenthal, David. 2005. *Consciousness and Mind*. New York: Oxford University Press.

Saxe, Rebecca. 2009. "Theory of Mind (Neural Basis)." In *Encyclopedia of Consciousness*, edited by William Banks. Boston: Elsevier/Academic Press.

Saxe, Rebecca. 2010. "The Right Temporo-Parietal Junction: A Specific Brain Region for Thinking about Thoughts." In *Handbook of Theory of Mind*, edited by Alan Leslie and Tamsin German. New York: Routledge.

Schrouff, Jessica, Vincent Perlbarg, Mélanie Boly, Guillaume Marrelec, Pierre Boveroux, Audrey Vanhaudenhuyse, Marie-Aurélie Bruno, Steven Laureys, Christophe Phillips, Mélanie Pélégrini-Issac, Pierre Maquet, Habib Benali. 2011. "Brain Functional Integration Decreases During Propofol-Induced Loss of Consciousness." *NeuroImage* 57 (1): 198-205.

Seager, William. 2004. "A Cold Look at HOT Theory." In *Higher-Order Theories of Consciousness: An Anthology*, edited by Rocco Gennaro, 255-275. Amsterdam: John Benjamins.

Tye, Michael. 2000. *Consciousness, Color, and Content*. Cambridge: The MIT Press.