

ARTICLE

Consciousness and the Ethics of Human Brain Organoid Research

Karola Kreitmair 

Department of Medical History and Bioethics, School of Medicine and Public Health, University of Wisconsin – Madison, Madison, Wisconsin, USA

Email: kreitmair@wisc.edu

Abstract

The possibility of consciousness in human brain organoids is sometimes viewed as determinative in terms of the moral status such entities possess, and, in turn, in terms of the research protections such entities are due. This commonsense view aligns with a prominent stance in neurology and neuroscience that consciousness admits of degrees. My paper outlines these views and provides an argument for why this picture of correlating degrees of consciousness with moral status and research protections is mistaken. I then provide an alternative account of the correlation between moral status and consciousness, and consider the epistemic ramifications for research protections of this account.

Keywords: human brain organoids; moral status; consciousness; research ethics; epistemology

As neuroscientific research involving human brain organoids (HBOs) progresses, commentators are concerned about the degree of research protections such entities are due.¹ Much is made of the possibility that HBOs may develop the capacity for consciousness, which is seen as relevant to the question of research protections.² If an entity is conscious or capable of being conscious so the thinking goes, then it deserves greater consideration in the research context than if it is not conscious or capable of being conscious. Knowing that an HBO is conscious or capable of being conscious puts us well on our way of knowing what kind of research protections it is due. While this stance is highly plausible, it fails to sufficiently account for both the epistemological challenges associated with consciousness and the way in which consciousness confers moral status. In this paper, I will show that what we can know about the consciousness of HBOs is insufficient to determine the kind of research protections that are ethically required.

My paper has four sections. Section 1 provides an overview of the relevant epistemological commitments of the study of consciousness. Section 2 outlines a commonsense approach to correlating consciousness, moral status, and research protections. Section 3 provides arguments against this commonsense view and identifies the ramifications of this view being false. Section 4 presents an alternative view and identifies the ramifications of this alternative view being correct.

1. Epistemological Considerations Regarding Consciousness

The term “consciousness” admits of several interpretations. In the neuroscience literature, for example, consciousness is sometimes defined as comprising wakefulness and awareness of environment and of self.³ While such a definition may be useful in clinical settings, it is too demanding for explorations around HBOs. Consciousness is sometimes also conceived of as being a matter of “cognitive

accessibility.”⁴ Such a conception is also too complex and demanding to capture the basic idea of consciousness that is relevant for our purposes. The most fundamental conception of consciousness identifies it as the phenomenon of subjective experience. When I am in a conscious state, there is “something it is like” for me to be in that state.⁵ This conception of consciousness is sometimes called “phenomenal consciousness” or “qualitative consciousness” or “qualia.” It is the least demanding conception of consciousness and thus appropriate for investigating the ethical impact of entities with inchoate consciousness at most. As such, it is the conception that will be employed in this paper.⁶

In the discussion around the ethics of HBO research, significant attention is focused on the ethical implications of HBOs that may develop consciousness or the capacity for consciousness. Julian Koplin and Julian Savluescu argue that if HBOs are conscious or “could plausibly be conscious,” stricter research protections would be required, including minimizing possible suffering and minimizing the number of HBOs used.⁷ Andrea Lavazza maintains that if HBOs “are endowed with some degree of consciousness [...] strong ethical objections [to research use] should not be ignored.”⁸ Lavazza and Marcello Massimini suggest that if HBOs were to develop a “primitive level of consciousness,” it would be appropriate to include them in the calculus of “hedonistic consequentialism.”⁹

Motivating the focus on consciousness in HBOs is the underlying assumption that consciousness provides some grounds for moral status.¹⁰ We view valanced experiences, such as pleasure and suffering, as objectively, intrinsically good or bad, at least in part in virtue of their phenomenal character. In turn, as the bearers of such value-laden states, entities capable of experiencing such states are themselves bearers of some value. This implies that entities capable of experiencing such states have some moral status.¹¹ It is important to note that the claim that consciousness is *sufficient* for some moral status does not imply that consciousness is *necessary* for some moral status. Indeed, patients in a permanent vegetative state (PVS) are generally thought to possess some moral status. While there is disagreement as to whether patients in PVS have full moral status,¹² medico-ethical norms maintain that such patients matter morally, despite their permanent lack of consciousness. The grounds for such mattering may include membership in the species of *Homo sapiens*,¹³ or a continued biological drive, or relational properties¹⁴. In any event, this suggests that consciousness is likely not necessary for moral status. In addition, the claim that consciousness is sufficient for *some* moral status also does not imply that consciousness is sufficient for *full* moral status. We tend to view non-human animals as likely possessing some moral status though not full moral status. The non-human animals to whom we accord this lesser degree of moral status tend to be those we believe have some conscious experience. In general, animals that we think lack consciousness, such as ants, are viewed as having less moral status than animals that we think have potentially complex conscious states, such as chimpanzees.

In this paper, I merely maintain that consciousness is sufficient for some moral status. I remain silent on whether it is necessary for moral status and whether it is sufficient for full moral status.¹⁵ As such, when we consider the relationship between moral status and consciousness in HBOs, we are focusing only on moral status conferred via consciousness. I will refer to this as “moral status-qua-consciousness.” This leaves open whether HBOs possess moral status that is grounded in sources besides consciousness.

With consciousness increasingly being discussed as a potential source of moral status for HBOs, it is important to ask how we could know if an HBO is conscious or is capable of being conscious. Some authors have explicitly taken up this challenge.¹⁶ This epistemic question is essential in understanding how the possibility of consciousness in an HBO might determine the sorts of research protection it should receive. How, indeed, could we know whether an HBO is conscious? How can we know that *any* entity is conscious? And, how can we know anything about the nature of the conscious state of an entity? The only entities for whom we can have direct knowledge of conscious states are ourselves.¹⁷ For all other entities, we must use inference to learn anything about their conscious state.¹⁸ In general, we use inference from behavior. In entities that exhibit verbal behavior, such as neurotypical, verbal, human adults, we ask them to provide us with verbal reports about their experience. Clinicians ask patients to rate their pain on a 10-point scale.¹⁹ In non-verbal entities, such as babies or non-human animals, we use non-verbal behaviors. If a baby winces and cries, we infer it is in pain.²⁰ Even in non-human animals, such as mice, we use behavioral coding systems that correlate the mouse behavior with subjective pain experiences.²¹

Of course, inference from behavior is not possible in non-behavioral entities. In non-behavioral humans, for instance, we cannot use behavior to determine the presence of conscious states. Consider, for example, patients misdiagnosed as being in the vegetative state²²—a state which by definition precludes consciousness—when in fact these patients retain “islands” of consciousness.²³ In such patients, the absence of behavior is mistakenly taken as evidence of the absence of consciousness.

Similarly, inference from behavior fails in alien entities for whom we lack the “dictionary” with which to decode the behavior. An alien arriving on Earth may exhibit very specific behaviors. These behaviors may even be correlated in a nomological fashion with specific conscious states. But since this correlation is alien, and we have no way of connecting behaviors and conscious states, we cannot infer anything about the alien conscious states (neither their presence nor their nature) from the alien behavior.

Consequently, when inference from behavior is not possible to determine the presence or nature of conscious states, we are left only with inference from whatever objectively observable phenomena we can access. Given that we are interested in consciousness, we look at brain states.²⁴ The process here is similarly one of inference. Consider two entities: Entity A is the brain of a non-behavioral entity and Entity B is the brain of a verbal, neurotypical adult human of whom we know that she is conscious. If Entity A’s brain is sufficiently similar to that of Entity B, and if we can identify those brain states in Entity B that are sufficient for consciousness, we can infer from the presence of those brain states in Entity A that A is also conscious.²⁵ Indeed, this is the process utilized by Adrian Owen et al. in detecting consciousness in a patient diagnosed as being in the vegetative state.²⁶ Owen and his colleagues conducted functional magnetic resonance imaging (fMRI) on an unresponsive patient who fulfilled all clinical criteria for the vegetative state diagnosis. On the fMRI, Owen et al. observed activation in areas of the cortex indistinguishable from those in verbal, neurotypical adults having conscious experiences.²⁷ From this, they concluded that this patient was conscious, despite an inability to manifest this consciousness behaviorally.

2. The Neural Correlates of Consciousness and the Degrees View of Consciousness

Making inferences about the presence and nature of consciousness on the basis of brain states is, of course, the central aim of the search for the *neural correlates of consciousness* (NCC). David Chalmers defines the neural correlates of consciousness as follows: “A [neural correlate of consciousness] 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.”²⁸ If we can identify the neural correlates of consciousness, it should be possible, at least in theory, to detect the presence of consciousness simply by detecting whether or not the brain is in a certain state. This could be very helpful in telling us about the presence of conscious states in non-behavioral entities, such as patients diagnosed as being in the vegetative state, for whom inference from behavior is not possible. As such, neuroscientists and philosophers have a long history of searching for the NCC.²⁹

Identifying the NCC is also crucial to determining whether an HBO is conscious. Given that HBOs are non-behavioral entities, the only method we have of detecting consciousness in an HBO is via inference from the neural correlates of consciousness. As noted earlier, this requires that the neural structure of an HBO be sufficiently similar to that of a neurotypical human brain. I have not specified what it means for two neural structures to be “sufficiently similar,” and it is not obvious how exactly this would be specified across different types of entities, such as the brains of human individuals and HBOs. For the purposes of this paper, it may suffice that such a similarity is possible, even if it cannot be precisely described. Prudential reasons suggest embracing a latitudinarian stance toward this similarity, so that we err on the side of inclusivity when it comes to detecting NCC.

Some may object that it is not accurate that HBOs are (or always will be) non-behavioral entities. Advances are being made in assembling multiple brain organoids into what are known as assembloids.³⁰ Moreover, given that the absence of blood flow to the organoid has been a major limiting factor to maturation, efforts are underway to vascularize brain organoids.³¹ In addition, photoreceptor cells have been connected with cerebral neurons in HBOs, making HBOs sensitive to the input of external light

stimuli.³² Finally, HBOs have been connected with explanted mouse spinal cords, resulting in HBO-caused contractions of adjacent muscle tissue.³³ This all suggests that it might eventually be possible for HBOs to exhibit what could be interpreted as “behavior.” However, while it is true that the HBO would then not be a non-behavioral entity, the behavior would be so unlike what we understand, that it would not permit inference on the basis of our own behavior. This kind of behavioral entity would be akin to the alien discussed above. We lack the dictionary to translate such an entity’s behaviors into conscious states.

Leading candidates for a theory of the NCC include the Integrated Information Theory (IIT), the Global Workspace Theory (GWT), and the Higher-Order Thought Theory (HOT). In this section I will focus on IIT, although I will return to the others later. IIT posits that a brain state is conscious if and only if it carries “integrated information,” which means that the effective informational content of the whole is greater than the informational content of the parts.³⁴ Integrated information is a scalar measure, denoted by ϕ . States that maximize ϕ are conscious.³⁵ IIT predicts that the NCC are located in the posterior hot zone, in the cerebral cortex.³⁶ In addition, IIT permits for an empirical measure of brain complexity, the Perturbational Complexity Index (PCI). The PCI is defined as the normalized Lempel–Ziv complexity of the spatiotemporal pattern of cortical activation, measured via encephalography (EEG), triggered by a direct transcranial magnetic stimulation (TMS) perturbation.³⁷ To compute the PCI of a brain state, it is sufficient to assess the EEG response to TMS. The PCI has been shown to discriminate reliably between conscious and non-conscious individuals in various disorders of consciousness.³⁸

The prevalent understanding of consciousness is informed by a conception of consciousness admitting of degrees. Likely originating in the clinical study of disorders of consciousness, consciousness is generally conceived of as existing on a single quantitative scale, with less conscious states toward one end of the scale and more conscious states toward the other. The extent to which a state is conscious is sometimes referred to as its “level of consciousness.”³⁹ Consciousness is described as “a scale ranging from total unconsciousness (e.g., death and coma) to vivid wakefulness,”⁴⁰ and as “a continuous variable.”⁴¹ Coma, general anesthesia, and the vegetative state (VS) are at the bottom of the scale, making those states “less conscious,” sleep and the minimally conscious state (MCS) are toward the middle, making those states “more conscious,” and wakefulness is at the top of the scale, making that state “most conscious.” Figure 1, which is reproduced from a 2013 article by Melanie Boly and colleagues,

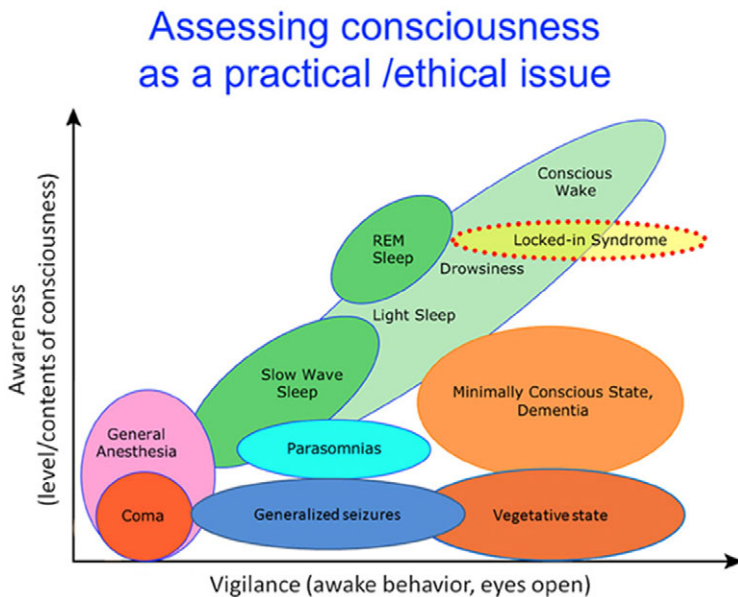


Figure 1. Levels of consciousness (reproduced from Boly et al.⁴³).

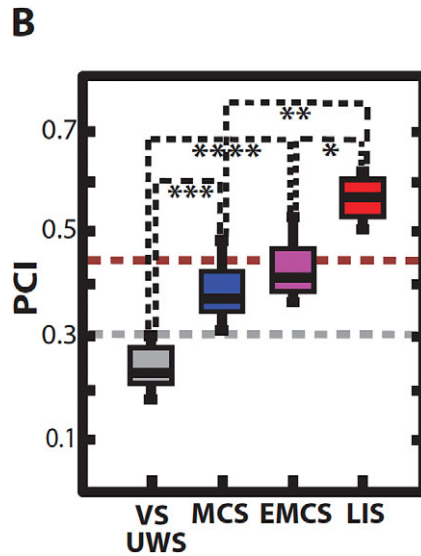


Figure 2. PCI discriminating between the “levels of consciousness” (reproduced from Casali et al.⁴⁶).

shows the ordering of conscious states along a single scalar dimension, with the y -axis representing the level of consciousness. Indeed, the introduction of MCS+ and MCS– diagnoses further distinguish between the levels of consciousness.⁴²

Indeed, some theories of NCC lend themselves directly to quantifying levels of consciousness. Since the PCI is a scalar measure, it can delineate between these “levels” of consciousness, with states at the lower end of the scale having a lower PCI and states at the higher end of the scale having a higher PCI.⁴⁴ Figure 2 shows the PCI computed from EEG measurements of patients who are unconscious in the vegetative state (PCI < 0.31), at “low degrees of consciousness” in MCS or emerging from MCS (PCI 0.32–0.49), and at “high degrees of consciousness” in the locked-in state (LIS) (PCI 0.51–0.62). As Adenauer Casali et al. state, the “PCI is sensitive to graded changes in the level of consciousness.”⁴⁵ The PCI, with its normalized values between zero and one, can quantify “how conscious” a brain state is.

3. A Neat Picture

The levels view of consciousness fits neatly with widely held intuitions about consciousness and moral status, and as a corollary, with intuitions about consciousness and research protections. Many believe that moral status admits of degrees.⁴⁷ While both a person and sentient non-human animals, such as a mouse, matter morally, the former is generally thought to have full moral status, whereas the latter only has partial moral status. This explains why, *ceteris paribus*, it constitutes a greater tragedy if a person is killed than if a mouse is killed. The difference in moral status is often attributed to a difference in cognitive capacities, with persons, but not non-persons, thought to possess rationality, self-awareness, and future-oriented plans.⁴⁸ Assuming that it is correct that moral status admits of degrees, and assuming that it is correct that consciousness is sufficient for some moral status, namely moral status-qua-consciousness, then it is plausible that entities that are more conscious have more moral status-qua-consciousness and entities that are less conscious have less moral status-qua-consciousness. This would support the intuition that, *ceteris paribus*, it is more morally unproblematic to withdraw life-sustaining treatments from a patient in PVS than from a patient in MCS.⁴⁹ In part, it seems, this is because the former lacks consciousness, whereas the latter has some consciousness, and because some moral status is bestowed in virtue of consciousness.⁵⁰

The view that moral status admits of degrees goes hand in hand with the view that entities with greater moral status are due greater research protections and entities with lesser moral status are due lesser

research protections. This explains why, *ceteris paribus*, it is thought to be permissible to use mice (who are generally thought to have only partial moral status) in research settings in which it would be impermissible to use persons (who are generally thought to have full moral status).

Putting this all together, we are left with a neat picture. If the level of consciousness an entity is capable of is correlated with the degree of moral status an entity has, and the degree of moral status an entity has is correlated with the level of research protections that are due that entity, then the level of consciousness an entity is capable of is correlated with the level of research protections that are due that entity. Moreover, given that we can empirically measure the level of consciousness of an entity via the PCI, we can empirically determine the level of research protections that are due to an entity.

Now, as I have stated, consciousness is only one possible source of moral status, and so the moral status that is conferred through consciousness may not exhaust the moral status of an entity. This explains why a patient in PVS is generally thought to possess a high degree of moral status, while a conscious non-human animal is generally thought to possess a lower degree of moral status. As such, it is possible that an HBO possesses moral status that is grounded in sources other than consciousness. However, it is plausible that, unlike with other entities, consciousness is the main source of moral status for an HBO. Research has been ongoing with human organoids of all sorts, including liver, kidney, prostate, and pancreas.⁵¹ With the exception of the gonadal organoids, the discussion around ethical issues related to these organoids focuses entirely on their provenance, in particular the issue of how to obtain consent from the stem cell donor.⁵² This suggests that such human non-brain organoids themselves are not seen as intrinsically morally valuable, but only extrinsically so. The exception of gonadal organoids to the above pattern is telling, because gonadal organoids may one day be used for reproductive purposes. As such, gonadal organoids raise similar ethical issues to human embryos, which are thought by many to possess (at least) partial moral status.⁵³ What this suggests is that the source of intrinsic moral status of HBOs stems from the fact that HBOs replicate the human organ that is the seat of cognition and consciousness. This suggests that the capacity for consciousness (perhaps along with the capacity for cognition) is the main source of moral status for an HBO.⁵⁴ As such, the observation holds that if we can empirically measure the level of consciousness in an HBO (e.g., via the PCI) and if this level is correlated with the degree of moral status-qua-consciousness, which in turn is correlated with the degree of research protections, then the empirical measure of consciousness informs us of the degree of research protections an HBO ought to receive.

4. An Alternative View

While the scenario painted in the previous section makes for a neat picture of how we can determine the type of research protections an HBO deserves, it is based on a fundamental mistake: It is incorrect that consciousness admits of degrees. Conscious states are not ordered along a single dimension of “more conscious” on one end, and “less conscious” on the other. The notion that consciousness admits of degrees or comes in levels is tempting. It makes sense to think of slowly waking up as slowly “turning up the dimmer” on consciousness. But on further reflection, it is clear that this is not how consciousness works. If there is a phenomenal experience, then that phenomenal experience has a determinate phenomenal content. Certainly, this content may be more or less varied, may be more or less bright or clear, may be more or less multi-sensory, and may be more or less mediated by concepts. But it makes no sense to suggest the content is more or less phenomenally determinate. There either is a phenomenal “what-it’s-likeness” or there is not. Andrew McKilliam sums this up by stating, “An entity either instantiates a determinate experience and is [...] conscious, or does not and is [...] not generically conscious.”⁵⁵ Alternatively, in the words of Bayne and colleagues, “A sighted person might be conscious of more than someone who is blind, but they are not more conscious than the blind person is.”⁵⁶

The ramification is clear. If there are no levels of consciousness, then the level of consciousness cannot be correlated with the degree of moral status (or degree of moral status-qua-consciousness). This suggests that we cannot infer the degree of research protections an HBO deserves from an empirical measurement of the “degree of consciousness.”

An alternative to the levels view and the view that I am advocating for is that moral status-qua-consciousness is differentiated not according to the “level of consciousness” but according to

the content of the conscious experience. In other words, moral status is conferred not in virtue of how conscious a state is, since all states are either conscious or not, but rather in virtue of what the phenomenal content of the state is, that is, what it is like to be in that state.⁵⁷ There are different properties that differentiate the contents of conscious states. Determining these properties is the subject of phenomenological exploration.⁵⁸ A thorough such exploration is beyond the scope of this paper. However, for our purposes, it is helpful to at least gesture at some such properties. A state may have phenomenal content from one, two, three, four, five, or more senses integrated in the phenomenal experience. It may be reflexive, in that the state may include the phenomenology of awareness of the conscious experience. It may include a conscious experience of the self as distinct from the world. This phenomenological experience of self may be rudimentary or sophisticated. A state may possess phenomenological diachronicity, that is, the sense of experiences existing over time. A phenomenological experience may be mediated by conceptual awareness. It may also be embedded in a narrative. Finally, the experience may be valenced, that is, there may be a “goodness” or “badness” to the phenomenological experience, such as we see in affective states of pleasure and suffering.

I posit that entities capable of more distinct and complex phenomenal states possess greater moral status-qua-consciousness than entities capable of fewer distinct and complex phenomenal states.⁵⁹ I will provide an example to elucidate this claim. Imagine an entity that is capable of experiencing either a light state or a dark state, and nothing more. This entity can be in three possible states: (1) unconscious, that is, there is nothing it is like to be this entity, (2) conscious and experiencing light, or (3) conscious and experiencing dark. Contrast this entity with an entity that has a thousand possible phenomenal states. Both entities have the capacity for consciousness. Both entities have the same “level” of consciousness. But, I argue, the latter entity possesses more moral status-qua-consciousness than the former. Note that this is not because the latter creature may be capable of behaving in more ways. That may or may not be true. But imagine that neither of these creatures exhibits any behavior. I maintain that if the latter creature were to die, it would be a greater loss than if the former were to die. This is because there is value and thus moral status in the experiencing or potential to experience itself. Moral status-qua-consciousness is bestowed simply in virtue of phenomenal experiences, that is, what it’s like to be that entity. There is no additional value to the entity’s moral status-qua-conscious when such experiences are veridical, useful, or anything else. Such considerations may feature in other sources of moral status, but not in moral status-qua-consciousness.

It is conceivable that with respect to moral status, some phenomenal states are more important than others.⁶⁰ Indeed, several commentators have argued that it is the potential for suffering that would make interventions in HBOs morally problematic.⁶¹ Perhaps pleasure and suffering, that is, valenced phenomenal states, are particularly relevant with respect to moral status-qua-consciousness? While this is a sensible suggestion, it may be premature to grant valenced states special status in this respect. Valenced phenomenal states may lose their luster or sting in the absence of other phenomenal features, such as phenomenal diachronicity.⁶² Is a negatively valenced phenomenal state more morally weighty than, for example, a blue phenomenal state, when neither state contains a phenomenal property of continuity through time? Perhaps, but this is far from obvious. Rather than apportioning valenced phenomenal states particular importance with respect to moral status-qua-consciousness, what this suggests is that determining moral status-qua-consciousness is not a simple task of summing up all possible phenomenal states of a given entity. Rather, it may be the case that determining such moral status must take into account more nuanced dependence relations between different types of phenomenal states (e.g., negative valence has added moral weight if there is potential for phenomenal diachronicity and no added moral weight if there is no potential for phenomenal diachronicity). It certainly seems plausible that certain phenomenal properties loom larger with respect to moral status-qua-consciousness, in virtue of their being enabling conditions for other, more complex phenomenal properties.

Let us return to our considerations of HBOs. If I am correct that moral status-qua-consciousness is correlated with the number and complexity of phenomenal states an entity is capable of, then one ramification is that the neat picture correlating empirical measurements to research protections in HBOs falls apart. If it is properties of the content of possible states of consciousness that confer moral status-qua-consciousness onto an HBO, then we need to have insight into that content to determine the moral status-qua-consciousness of the HBO. As argued, the only way to have insight into the conscious states of

non-behavioral entities such as an HBO is via NCC. However, it is not clear that the NCC can provide such fine-grained insight into the conscious states of an entity.

IIT, as discussed above, generates the PCI as an empirical measure. The PCI is a normalized value between zero and one. While such a measure is well-suited to provide scalar information, it does not give insight into the content of the experiences that give rise to the PCI value. Indeed, the PCI is designed to abstract away from complexity and condense brain patterns into a single value. But of course, there are other candidate theories for NCC.⁶³ However, it is not clear that any NCC can provide insight into the content of phenomenal experience in a way that would be sufficient to provide us with the information required to make judgments about the moral status-qua-consciousness. In addition to IIT, a leading contender for NCC is the global workspace theory (GWT). GWT holds that a state is conscious if and only if it is present in the global neuronal workspace. This makes the state *accessible* to multiple systems including memory, attention, and perception.⁶⁴ GWT predicts that certain pyramidal cells found in the prefrontal, cingulate, and parietal regions, along with associated thalamocortical loops, form a “neuronal workspace.”⁶⁵ Conscious content, then, is encoded by the activation of a subset of these neurons.⁶⁶ While GWT provides testable predictions about which brain regions are active during conscious experience, it does not provide predictions so fine-grained as to be able to read off the precise content of the phenomenal state. A further theory of the NCC, the higher order thought theory (HOT) posits that a state is conscious if and only if one can represent oneself as being in such a state.⁶⁷ This suggests consciousness requires activity in the prefrontal cortex. While it is certainly testable whether there is activation in the prefrontal cortex during a conscious experience, it is not clear what information this provides about the specific content of the phenomenal experience.

In sum, while it is conceivable that we may one day have a way of empirically detecting the specific content of the conscious experience of a non-behavioral entity, this does not seem likely. Without an empirical means of gaining insight into the content attributes of conscious states that an HBO is capable of, we will not be able to determine the moral status-qua-consciousness conferred upon that HBO in virtue of such conscious states (or the capacity for such conscious states). This means we cannot rely on considerations of the moral status-qua-consciousness of an HBO to determine appropriate research protections. The upshot is that we need some other way, independent of considerations of consciousness, to determine what sorts of research protections are appropriate for an HBO. This likely involves a value judgment about the amount of tolerance we ought to have in conducting research for getting the moral status of the research subject wrong. Further exploration of this problem is warranted.

In conclusion, in this paper, I present an argument for why the NCC are the only way of gaining insight into the presence and nature of conscious states of an HBO. I consider a commonsense approach to correlating “levels” of consciousness with moral status and, by extension, with research protections. I show how empirical measurements generated by certain NCC fit neatly with this commonsense view. I then present an argument for why this commonsense view is mistaken and offer an alternative approach. Finally, I consider the ramifications of the failure of the commonsense view, including that it is unlikely that what we can know about the conscious states of HBOs will ever be sufficient to provide insight into the type of research protections they are due.

Notes

1. Greely HT. Human brain surrogates research: The onrushing ethical dilemma. *The American Journal of Bioethics* 2021;**21**(2):34–45.
2. Koplin JJ, Savulescu, J. Moral limits of brain organoid research. *Journal of Law, Medicine & Ethics* 2019;**47**(4):760–67.
3. Calabrò RS, Cacciola A, Bramanti P, Milardi D. Neural correlates of consciousness: What we know and what we have to learn! *Neurological Sciences* 2015;**36**(4):505–13.
4. Block NJ. On a confusion about a function of consciousness. *Brain and Behavioral Sciences* 1995;**18**(2):227–47.
5. Nagel T. What is it like to be a bat? *The Philosophical Review* 1974;**83**(4):435–50.

6. The term ‘sentience’ is sometimes used as distinct from ‘consciousness’. It is used, for instance, in discussion around pain or heat perception. See: DeGrazia D. Sentience and consciousness as bases for attributing interests and moral status: Considering the evidence and speculating slightly beyond. In Johnson SM, Fenton A, Shriver A, eds. *Neuroethics and Nonhuman Animals, Advances in Neuroethics*. Cham: Springer International Publishing; 2020:17–31. However, the way I am using consciousness in this paper encompasses all phenomenal experiences and so also encompasses any states that are otherwise described as ‘sentient’.
7. See [note 2](#), Koplin, Savulescu 2019.
8. Lavazza A. Potential ethical problems with human cerebral organoids: Consciousness and moral status of future brains in a dish. *Brain Research* 2021;**1750**:147146.
9. Lavazza A, Massimini M. Cerebral organoids: Ethical issues and consciousness assessment. *Journal of Medical Ethics* 2018;**44**(9):606–10.
10. Shephard J, Levy N. Consciousness and morality. In *Oxford Handbook of the Philosophy of Consciousness*. Oxford, UK: Oxford University Press; 2020:654–72.
11. See [note 10](#), Shephard, Levy 2020, at 657.
12. Syd Johnson, for example, argues that PVS patients have full moral personhood. See: Johnson LSM, *The Ethics of Uncertainty: Entangled Ethical and Epistemic Risks in Disorders of Consciousness*, New York, NY, USA: Oxford University Press; 2021. On the other hand, Grant Gillett, argues that PVS patients lack full moral status. See: Grant Gillett G, *Consciousness, the brain and what matters*. *Bioethics* 1990;**4**(3):181–98.
13. Ronald D, *Life’s Dominion*, New York, NY: Penguin Random House; 1994.
14. Kittay EF. Equality, dignity, and disability. In Lyons MA, Waldron F, eds. *Perspectives on Equality The Second Seamus Heaney Lectures*. Dublin: The Liffey Press; 2005.
15. My sense is that consciousness is both not necessary for moral status and not sufficient for full moral status, in part because of our intuitions concerning permanently unconscious human individuals and concerning conscious non-human animals.
16. See [note 9](#), Lavazza, Massimini 2018; Shepherd J. Ethical (and Epistemological) issues regarding consciousness in cerebral organoids. *Journal of Medical Ethics* 2018;**44**(9):611–12; Niikawa T, Hayashi Y, Shepherd J, Sawai T. Human brain organoids and consciousness. *Neuroethics* 2022;**15**(1):5.
17. Descartes R. *The Philosophical Writings of Descartes: Volume 2*. Cambridge: Cambridge University Press; 1984.
18. This is of course the Problem of Other Minds, which has received much attention from epistemologists since at least the time of Descartes. David Chalmers has argued that inference to the best explanation is likely the best “solution to the problem of other minds [...] we are going to get.” See: Chalmers DJ. *The Conscious Mind*. Oxford, UK: Oxford University Press; 1996:246.
19. “Your Pain on a Scale of 1–10? Check Out a New DOD Way to Evaluate Pain,” Military Health System, October 13, 2022, <https://www.health.mil/News/Articles/2022/10/13/DVPRS-pain-scale>.
20. Williams ACDC. Facial expression of pain: An evolutionary account. *Behavioral and Brain Science* 2002;**25**(4):439–55.
21. Langford DJ, Bailey AL, Chanda ML, Clarke SE, Drummond TE, Echols S, Glick S, Ingrao J, Klassen-Ross T, Lacroix-Fralish ML, Matsumiya L, Sorge RE, Sotocinal SG, Tabaka JM, Wong D, van den Maagdenberg AMJM, Ferrari MD, Craig KD, Mogil JS. Coding of facial expressions of pain in the laboratory mouse. *Nature Methods* 2010;**7**(6):447–49.
22. Fins JJ, Master MG, Gerber LM, Giacino JT. The minimally conscious state: A diagnosis in search of an epidemiology. *Archives of Neurology* 2007;**64**(10):1400–05.
23. Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD. Detecting awareness in the vegetative state. *Science* 2006;**313**(5792):1402.
24. Of course, it is possible to be skeptical of the notion that we can ‘know’ anything about minds that are not our own, whether such knowledge is based on inference from behavior or inference from brain states. But such radical skepticism jeopardizes virtually all scientific approaches to the mind, so we will bracket it for the purposes of this paper.

25. It is true that in order for this inference to work, we must assume that conscious states are necessarily correlated with brain states in a nomological fashion. However, as stated above, without such a bedrock assumption, we are doomed to skepticism.
26. See [note 23](#), Owen et al. 2006.
27. See [note 23](#), Owen et al. 2006.
28. Chalmers DJ. What is a neural correlate of consciousness? In Metzinger T, eds. *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press; 2000:17–40.
29. Dehaene S, Naccache L. Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition* 2001;**79**(1–2):1–37; Boly M, Massimini M, Tsuchiya N, Postle BR, Koch C, Tononi G. Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. *Journal of Neuroscience* 2017;**37**(40):9603–13; Koch C, Massimini M, Boly M, Tononi G. Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience* 2016;**17**(5):307–21; Nani A, Manuella J, Mancuso L, Liloia D, Costa T, Cauda F. The neural correlates of consciousness and attention: Two sister processes of the brain. *Frontiers in Neuroscience* 2019;**13**:1169.
30. Schmidt C. The rise of the assembloid. *Nature* 2021;**597**(7878):S22–23.
31. Matsui TK, Tsuru Y, Hasegawa K, Kuwako K-I. Vascularization of human brain organoids. *STEM CELLS* 2021;**39**(8):1017–24.
32. Quadrato G, Nguyen T, Macosko EZ, Sherwood JL, Yang SM, Berger DR, Maria N, Scholvin J, Goldman M, Kinney JP, Boyden ES, Lichtman JW, Williams ZM, McCarroll SA, Arlotta P. Cell diversity and network dynamics in photosensitive human brain organoids. *Nature* 2017;**545**(7652):48–53.
33. Giandomenico SL, Mierau SB, Gibbons GM, Enger L, Masullo L, Sit, T, Sutcliffe M, Boulanger J, Tripodi M, Derivery E, Paulsen O, Lakatos A, Lancaster MA. Cerebral organoids at the air-liquid interface generate diverse nerve tracts with functional output. *Nature Neuroscience* 2019;**22**(4):669–79.
34. Tononi G, Boly M, Massimini M, Koch C. Integrated information theory: From consciousness to its physical substrate. *Nature Reviews Neuroscience* 2016;**17**(7):450–61.
35. Jeziorski J, Brandt R, Evans JH, Campana W, Kalichman M, Thompson E, Goldstein L, Koch C, Muotri AR. Brain organoids, consciousness, ethics and moral status. *Seminars in Cell & Developmental Biology* March 23, 2022.
36. See [note 29](#), Koch et al. 2016.
37. Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Casali KR, Casarotto S, Bruno M-A, Laureys S, Tononi G, Massimini M. A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine* 2013;**5**(198):198ra105–198ra105.
38. Casarotto S, Comanducci A, Rosanova M, Sarasso S, Fecchio M, Napolitani M, Pigorini A, Casali AG, Trimarchi PD, Boly M, Gosseries O, Bodart O, Curto F, Landi C, Mariotti M, Devalle G, Laureys S, Tononi G, Massimini M. Stratification of unresponsive patients by an independently validated index of brain complexity. *Annals of Neurology* 2016;**80**(5):718–29.
39. Boly M, Seth AK, Wilke M, Ingmundson P, Baars B, Laureys S, Edelman DB, Tsuchiya N. Consciousness in humans and non-human animals: Recent advances and future directions. *Frontiers in Psychology* 2013;**4**:625.
40. Seth AK, Dienes Z, Cleeremans A, Overgaard M, Pessoa L. Measuring consciousness: Relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences* 2008;**12**(8):314.
41. Dehaene S, Changeux JP. Neural mechanisms for access to consciousness. In *The Cognitive Neurosciences*, 3rd Ed. Cambridge, MA: Boston Review; 2004:1145–57.
42. Bayne T, Hohwy J, Owen AM. Are there levels of consciousness? *Trends in Cognitive Sciences* 2016;**20**(6):405–13.
43. See [note 39](#), Boly et al. 2013.
44. See [note 37](#), Casali et al. 2013 and [note 38](#), Casarotto et al. 2016.
45. See [note 37](#), Casali et al. 2013.
46. See [note 37](#), Casali et al. 2013.
47. DeGrazia D. Moral status as a matter of degree? *The Southern Journal of Philosophy* 2008;**46**(2):181–98.

48. Kant I, Gregor M, trans. *Groundwork of the Metaphysics of Morals*. Cambridge Texts in the History of Philosophy. Cambridge, U.K.; New York: Cambridge University Press 1998; Tooley T. Abortion and infanticide. *Philosophy & Public Affairs* 1972;**2**(1):37–65.
49. Demertzi A, Ledoux D, Bruno M-A, Vanhaudenhuyse A, Gosseries O, Soddu A, Schnakers C, Moonen G, Laureys S. Attitudes towards end-of-life issues in disorders of consciousness: A European survey. *Journal of Neurology* 2011;**258**(6):1058–65.
50. See Guy Kahane and Savulescu on this issue. Kahane G, Savulescu JJ. Brain damage and the moral significance of consciousness. *The Journal of Medicine and Philosophy* 2009;**34**(1):6–26.
51. Mollaki V. Ethical challenges in organoid use. *BioTech* 2021;**10**(3):12.
52. See [note 51](#), Mollaki 2021.
53. Brown MT. The moral status of the human embryo. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine* 2018;**43**(2):132–58.
54. See [note 50](#), Kahane, Savulescu 2009.
55. McKilliam AK. What is a global state of consciousness? *Philosophy and the Mind Sciences* 2020;**1**(II): p.4.
56. See [note 42](#), Bayne, Hohwy, Owen 2016.
57. Shepherd J. *Consciousness and Moral Status*. London: Taylor & Francis; 2018.
58. Husserl E. *Ideas: General Introduction to Pure Phenomenology*. Milton Park, UK: Routledge; 2012; Merleau-Ponty M, Landes D trans. *Phenomenology of Perception*. Milton Park, UK: Routledge; 2013.
59. My view here is similar to that of Shepherd in *Consciousness and Moral Status*.
60. I thank an anonymous reviewer for this suggestion. Julian Savulescu raised a similar point in personal communication.
61. See [note 9](#), Lavazza, Massimini 2018.
62. For the distinction between diachronic and synchronic unity, see Rashbrook O. Diachronic and synchronic unity. *Philosophical Studies* 2013;**164**(2):465–84.
63. IIT may also be able to generate other types of measurements.
64. See [note 29](#), Dahaene, Naccache 2001.
65. See [note 41](#), Dahaene, Changeux 2004.
66. See [note 41](#), Dahaene, Changeux 2004.
67. Lau H, Rosenthal D. Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences* 2011;**15**(8):365–73.