

## Ethical Implications of Indeterminate Cases of Consciousness

It's often of great ethical importance whether a particular agent, or a particular class of agents are conscious. For example, animals which are conscious need to be covered by animal welfare legislation. And different philosophical and scientific accounts of consciousness have importantly different implications about which creatures are conscious. For example, so-called higher-order accounts of consciousness (Carruthers 2016) have been said (Seager 2004) to imply that only human beings (and perhaps great apes) are ever conscious. In this talk I'll explore the ethical implications of the following possibility: there are real phenomena in the brain which fit distinct and incompatible philosophical or scientific accounts of what consciousness is; further there's no fact of the matter which of these phenomena counts as being consciousness, since each of the phenomena fits our pre-theoretic conception of consciousness equally well. I'll first motivate the idea that this's a realistic possibility. Then I'll discuss some problems about ethical decision making which arises if the possibility is realized. I'll focus on animal welfare, though there are cases involving coma patients where related issues arise. Consider a case where distinct accounts of consciousness imply conflicting answers as to whether, say, crustaceans are ever conscious. Plausibly, if crustaceans *are* conscious, then we have at least prima facie reason to worry about causing them pain, whilst if they are not conscious, we have no such reasons. But what if they undergo mental states which count as conscious under some of the reasonable competing philosophical and cognitive scientific accounts of consciousness but not by others, and some of those mental states are otherwise relevantly similar to conscious pains, except insofar as they meet some but not others of the criteria? Given the assumption that there's no fact of the matter about whether these mental states are in fact conscious, and that being conscious seems to be a necessary condition on a mental state's bearing value/disvalue (or at least on it bearing one kind of value/disvalue), are we forced to concede that there's no fact of the matter about whether we have prima facie reasons to avoid inflicting pain-like states on crustaceans?

In the first part of the talk, I will briefly motivate the no fact of the matter claim itself, by motivating a claim I call '**Theory Equality in the Metaphysics of Consciousness**' (TEMIC), and arguing that the no fact of the matter claim follows from it:

**TEMIC:** Amongst the standard accounts of the metaphysics of 'consciousness', there are multiple, competing and inconsistent accounts of what consciousness really is, and none of these accounts are better than any of the others, because they all do equally well at preserving the (wide) conceptual role of CONSCIOUSNESS; it is therefore semantically indeterminate whether mental states which count as conscious according to some but not all of these theories count as cases of consciousness.

I'll motivate this by giving some reasons to think that higher-order theories of consciousness do better than first-order representationalist theories at capturing the theoretical claims to which ordinary thought about consciousness is committed, but worse at capturing the pattern of applications of the term in actual cases which ordinary speakers display. And I'll argue that, given this, it's at least not obvious that higher-order theories do either better or worse than first-order representationalist theories, at giving us a 'real definition' of consciousness, and that if this is so, some kind of semantic indeterminacy involving consciousness will result (at least if we favour semantic over epistemic theories of vagueness.) I'll also sketch some further ways in which TEMIC might come out true, even if this particular argument fails, and briefly state why I think it's far from obvious that none of these possibilities are realized.

Having motivated TEMC, I'll then go on to discuss the ethical implications if TEMC is true. Plausibly, pains are bearers of disvalue, because they are conscious mental states which are aversive to their subjects. But what if there are mental states which, whilst otherwise reasonable candidates for being pains, count as consciousness by some reasonable theories but not others, and so are indeterminate cases of consciousness. Does this mean that it is also indeterminate whether such mental states are bearers of moral disvalue, and therefore indeterminate exactly what duties we have to relieve or prevent them?

I'll first show that if the particular line of thought I gave motivating TEMC is correct, this isn't a merely theoretical question about some very rare edge cases. Rather, if it's indeterminate whether higher-order or first-order representationalist theories of consciousness are correct, then it's likely that a dramatically large number of animal species will be such that whether or not they count as capable of undergoing conscious mental states depends on whether or not we take a higher-order or first-order theory to be correct. This shows that it's important whether TEMC is true, since its truth threatens to bring widespread indeterminacy about value.

Then I'll go on to discuss whether holding that there are some mental states which count as indeterminate cases of consciousness, forces one to say that it's also indeterminate whether those mental states meet the enabling condition on being a bearer of value which being conscious provides. In particular, taking pain as a test case, since it's a clear case of a (dis)valuable conscious mental state, I'll discuss whether, in the case of mental states which are borderline cases of consciousness, but otherwise clear cases of pain, we are forced to say that, since there's no fact of the matter about whether they are conscious, there is no fact of the matter about whether or not they are bearers of disvalue, and hence no fact of the matter about what duties we have to prevent or relieve them. I'll first argue that if there is no way to treat the borderline cases of consciousness created by TEMC as cases where an enabling condition on possessing value is *partially* met, then we will have to say that there is no fact of the matter about whether or not pain-like mental states on the borderline are bearers of disvalue. The argument here is that if there is no way to see those mental states as meeting the condition for being bearers of (dis)value less than clearly conscious mental states, but more than unconscious ones, then there is no reason to assign them an ability to bear (dis)value to a degree intermediate between that of clearly conscious and clearly unconscious mental states, but that it would be arbitrary to assign them the same value-bearing ability as either the former or the latter. Hence, the only option left is to say that it's indeterminate whether they are bearers of value. I'll then discuss the prospects for making sense of the idea that pain-like states in the indeterminate zone partially but not fully meet the enabling criteria for being bearers of value that consists in being conscious. In particular, I'll briefly explore a model for such a view on which each such mental state gets assigned a real number greater than 0 and less than 1, representing the degree to which it meets the condition of being 'conscious', a number representing the (dis)value it would have, were it a clear case of a conscious mental state, and we then treat its (dis)value as the result of multiplying the second number by the first. I'll discuss some problems for this model, relating both to whether we can make sense of being a precise distance for being fully conscious in indeterminate cases, and whether we can make sense of the value an indeterminate mental state would have if fully conscious in the way the model requires. I won't reach any firm conclusions in such a short talk about whether the problem can be overcome, but I will draw some connections to the literature on the relationship between vagueness, and the notion of being true to a degree.

## References

Carruthers, Peter 2016: 'Higher-Order Theories of Consciousness', *Stanford Encyclopedia of Philosophy*, available at <https://plato.stanford.edu/entries/consciousness-higher/>

Seager, W. 2004: 'A Cold Look at HOT Theory', in Gennaro, R., (ed.) *Higher-Order Theories of Consciousness*, Philadelphia: John Benjamins.