PENULTIMATE DRAFT: forthcoming in *Rethinking Moral Status*, OUP, eds. Stephen Clarke, Hazem Zohny & Julian Savulescu

The moral status of conscious subjects

Joshua Shepherd[[1]](#footnote-1)

Keywords: phenomenal consciousness; moral status; cognitive sophistication; valence; intrinsic value; precautionary principle

**1 Theorizing about moral status**

Any account of moral status must specify the grounds of moral status. How are we to do that? Abstractly, two approaches are available.

A *members-first* approach begins by collecting judgments regarding who has moral status, and perhaps by collecting comparative judgments regarding who has higher degrees of moral status. Any account that begins with the judgment that humans have full (or the highest level of) moral status, and then seeks to justify that judgment, embodies this approach. So too does any account that begins with the judgment that healthy adult humans are at least the paradigmatic case of an entity with moral status.

This approach can seem epistemically modest, in the sense that it allows us to move from something we seem to be in a decent position to know, namely the grounds of the moral status of adult human beings, to elements that are more difficult to know, namely the grounds of whatever moral status other beings have. But I find the approach pernicious.

One reason is that it sets the theorist off on an unguided quest to discover the grounds of healthy adult human moral status, armed only with the intuition that it exists. But it turns out that this quest generates as little agreement as many other philosophical quests. Adult humans are complicated creatures, with a range of potentially morally relevant capacities and properties. Theorists have variously seized on many of these to offer accounts of the grounds of moral status. These include possession of self-consciousness (Tooley 1972), possession of sophisticated psychological capacities (McMahan 2002), possession of ‘typical human capacities’ (DiSilvestro 2010), possession of the capacity to participate in a ‘person-rearing relationship’ (Jaworska and Tennenbaum 2014), possession of a capacity for intentional agency (Sebo 2017), the ability to take oneself to be an end rather than a mere means in the sense that one can experience and pursue what is good for one (Korsgaard 2013), the capacity to suffer (Bentham 1996), possession of the genetic basis for moral agency (Liao 2010), and no doubt more. Of course some of these are friendlier to entities outside the tight circle of healthy adult humans, and some are not. What is striking, however, about many of these accounts is that they do not seek to justify the assumption that healthy adult humans are the paradigm case. Rather, this assumption justifies their search for the features in virtue of which the assumption must be true. But if healthy adult humans are not the paradigm case, the search may be headed in the wrong direction from the get go.

A second approach to the grounds of moral status is available. On a *value-first* approach, we first seek to isolate and theorize about the sources of value in the lives of entities who may have moral status. We then map whatever judgments regarding value we collect to considerations of moral status.[[2]](#footnote-2)

In the next section I articulate a fragment of such an account. I articulate a theory of what is valuable about *mental* life. (I do not assume that the value present in the lives of minded individuals qua minded individuals is the only value that could ground moral status: thus the fragmentary nature of the account.) In sections 3 and 4 I then seek to map this account of value onto considerations of moral status.

**2 Phenomenal consciousness and value**

I begin with an account of what is non-derivately valuable within a mental life. At the center of the account is phenomenal consciousness. This is a property, or really a set of properties, that attach to mental states and events. These are the properties of mental states or events in virtue of which there is ‘something it is like’ for the subject to token the state or event. If a mental state has phenomenal properties, then there is something it is like for the subject to be in that mental state. If a state lacks phenomenal properties, then it is unconscious in the sense that there is nothing it is like for the subject to be in that state.

Many share the intuition that phenomenal consciousness is somehow at the center of what is valuable about subjectivity. Without phenomenal consciousness, many think, being awake, being alive, being a subject, would be little different than being in coma, being dead, or lacking subjectivity altogether.

But what about phenomenal consciousness supports this intuition about value? Phenomenal consciousness is variegated. A conscious experience might contain, at some time or over some window of time, aspects grounded in sensory modalities (vision, olfaction, audition, proprioception), aspects grounded in cognitive capacities (memory, attention, inner speech, intention formation), aspects grounded in background levels of arousal (as when one feels sharp, or groggy), aspects grounded in emotions and moods. The complexity of conscious experience raises difficult questions. Are all aspects of conscious experience non-derivatively valuable, or only some? Are some aspects of conscious experience more valuable than others?

Very few philosophers have said much about these more specific questions (but see Siewert 1998, Kahane and Savulescu 2009, Kriegel 2019, Lee 2019). My account of what is non-derivatively valuable about consciousness attempts to begin to rectify this. But only to begin: I think a significant period of discussion and reflection is required.

My account of what we might call *phenomenal value* centers around a few ideas. The first can be put as follows.

[Affective-Evaluative Claim] It is necessary and sufficient for the presence of some (non-derivative) value in a conscious experience that the experience has evaluative phenomenal properties that essentially contain affective phenomenal properties. (Shepherd 2018a, 31)

The use of the term ‘affective phenomenal properties’ is ostensive. I am attempting to single out a class of phenomenal properties. Given that any proposed taxonomy of phenomenal properties is likely to run into controversy, it is perhaps easiest to do this by example. The kinds of properties I have in mind are these: the painfulness of pain, the pleasantness of pleasure, the distinct gut-located unease one might feel during a bout of food poisoning, the warm gut-located hum one might feel when thinking of one’s first love, the quickening of one’s attention coupled with the vibrant urge to react in some way when suddenly frightened by an unsuspected threat, the gnawing of boredom, the bubbling rise of frustration associated with an unsolvable puzzle. The category is broad and complex. What binds it together is the fact that the phenomenal properties in question are valenced in some way, along some dimension or other (or perhaps, in some cases, along multiple dimensions simultaneously).

Why do I have language about evaluative properties that essentially contain affective properties? Again, I am attempting to denote the relevant class. I wish to exclude one reading of ‘evaluative properties’ on which a conscious judgment that something is good or bad bears non-derivative value (since a conscious judgment, considered as a piece of cognitive phenomenology, may be thought to have evaluative properties, namely, those properties required to classify the judgment’s evaluative content).

The [Affective-Evaluative Claim] states that [a] without this phenomenologically recognizable affective valence, a conscious experience would not be non-derivatively valuable, and [b] with this affective valence, a conscious experience bears at least some non-derivative value. I say a few words about implications of this claim below.

The second idea behind my account of phenomenal value is this.

[Strong Evaluative Claim] It is necessary and sufficient for the presence of some non-derivative value in a subject’s mental life that the mental life contain episodes with essentially affective evaluative phenomenal properties. (35)

To state the same point from a different angle: a mental life with no affective experiences is a mental life without non-derivative value. As I put it elsewhere, ‘some mental item’s being non-derivatively valuable requires not just essentially affective evaluative properties, but phenomenal versions of these properties’ (37). My rationale for thinking this is two-fold. First, cases that strip this aspect of phenomenal consciousness away seem at the same time to strip away non-derivative value. Andrew Lee (2019) articulates a similar view, and attributes something like it to Jonathan Glover (2006). Lee calls this a neutral view on the intrinsic value of consciousness, according to which [a] consciousness itself is not intrinsically valuable, [b] some conscious states are intrinsically valuable, and [c] the intrinsically valuable states are so in virtue of their phenomenal character. This view is probably not identical to mine – differences may arise regarding the in-virtue-of relation that holds between phenomenal character and value, regarding the kind of value at issue (intrinsic or non-derivative), and regarding the particular states deemed valuable. But the view is obviously close. Lee does not argue in full for the neutral view, but he gives a case that gives one reason to accept it.

Consider two worlds that are empty save for a single creature inhabiting each world. In the first world, the creature has a maximally simple conscious experience that lacks any valence. Perhaps, for example, the creature has an experience of slight brightness. The creature’s experience is exhausted by this sparse phenomenology. In the second world, the creature is not conscious at all. (Lee 2019, 663)

If one has the intuition that the world with the conscious experience is no more valuable than the world without, one leans towards a neutral view. I do.

A second reason for accepting [Strong Evaluative Claim] is that, arguably, the nature of phenomenal consciousness is such that our phenomenally conscious states are present to us in a unique way. There is a directness in the relationship between a subject and the phenomenal properties of her experiences, at least in one sense. For without phenomenal consciousness, the relevant properties would not be present to her in the same way. In the case of affective experiences, this presence has consequences for the theory of value. For in the case of these experiences, what is present to the subject are items of non-derivative value, and it seems that the fact that they are present is crucial to the value that they bear. Non-conscious versions of these mental states – non-conscious emotional episodes, perhaps – may play important functional roles for the subject. But non-conscious versions do not bear non-derivative value.

I can explicate additional key ideas behind my account by way of answers to several questions, as follows.

First, what bears non-derivative value: phenomenal properties or conscious experiences? Another way of asking this question: do phenomenal properties only bear value individually, or can combinations of phenomenal properties bear value?

I think that combinations of phenomenal properties can bear value. To deny this would be to deny, for example, that the feeling of suffering attached to watching a loved one suffer is no less disvaluable than a similar feeling of suffering attached to watching a movie character suffer. I think that some ways phenomenal properties combine produce far more value than individual phenomenal properties could.

Second, how do phenomenal properties combine in ways relevant to the value that they bear? This is a complicated issue. Here I turn to work on the mereology of conscious experiences. And I follow theorists like Timothy Bayne (2010) and like Christopher Hill (2014) who point out a wide range of ways that multiple properties can be unified into a single experience. Importantly, some of the unity relations that bring disparate properties together into experiences are empirically tractable. In some cases subjects have the capacity to integrate information, and the capacity for integration helps explain how an experience can be more or less complex while remaining coherent.

So, consider an experience with affective phenomenal properties and many other properties besides. We have a further question to ask. In virtue of what does that experience bear value – in virtue of the affective properties alone, or in virtue of all of the properties constitutive of the experience?

I answer that the experience bears value in virtue of all of the properties that make it what it is. This adds a layer of complexity to the neutral view I endorsed above. For on this view, a non-affective property does not bear value in isolation. But as a part of a complex experience, a non-affective property can make important contributions to the value the experience bears.

Fourth, if an experience bears value in virtue of all of its properties, how do you determine how much value an experience bears?

In order to illuminate the ways experiences – mental episodes with more or less complex structures and shapes – bear value, I introduce the notion of an evaluative space. A subject’s evaluative space is a function of the interaction of sets of capacities – the subject’s affective-evaluative capacities, and the subject’s capacities for tokening evaluable elements. The latter class of capacities is broad, and includes perceptual and sensory registration capacities as well as capacities for thought and experience-generation and maintenance – imagination, deliberation, memory, language, inhibitory control are all obviously relevant. The former class of capacities involves the capacity to discriminate between the relevant evaluable materials, to generate experiences that ‘color’ these materials as evaluated on various dimensions, and the range of ways the subject can do so (i.e., with more or less intensity, with more or less vividness, with a greater or lesser number of modes of evaluation).

The subject’s evaluative space, then, depends in obvious ways upon her entire psychological architecture and her toolkit of perceptual, cognitive, agentive, and evaluative capacities. But the main emphasis here is upon the ways that her affective-evaluative experience generating capacities use and interact with the rest of her psychological systems. Think, by analogy, of a painter in a room. What she can produce depends in large part upon her capacities to paint. But it depends as well on the colors and brushes and canvasses she has available in the room. The colors, brushes, and canvasses are analogous to the psychological architecture generally. The skill at painting – what she can do with what she has – is analogous to the affective-evaluative experience-generating capacities.

This is just to introduce the notion of an evaluative space. What is the relationship of such a space to non-derivative value? I suggest that the *phenomenal richness* of a subject’s evaluative space tempers her capacity for undergoing highly valuable and disvaluable experiences. As I envision it, there is at least a rough correspondence between richness and capacity for (dis)value.

Phenomenal richness can be further understood in terms of three interacting metrics. First, the size of a subject’s evaluative space is measured by the number of properties and property-types a subject is capable of tokening (across different experiences). Second, the complexity of a subject’s evaluative space is measured by the number of properties and property-types she can bring together under a legitimate unity relation – that she can token, that is, within a single experience. Third, the coherence of a subject’s evaluative space is measured in terms of a relation of sense-making that holds between experiences. Coherence here holds over sets of experiences, and lends, among other things, a sense of structure and narrative to the flow of a subject’s experiences over time.

This is all abstract, and it takes a lot of work in sorting through different kinds of experiences to see why increases in size, complexity, and coherence might be thought to track differences in value. Here I wish to make only one further point about this idea. It is similar in ways to what many practical ethicists have said about the relation between cognitive sophistication and moral status. Here, for example, is David DeGrazia, reflecting on ways to justify giving unequal consideration to some animals (like humans) over others (like mice):

The most plausible specification . . . of the Unequal Consideration Model is a gradualist one, a sliding‐scale model, according to which sentient nonpersons deserve consideration in proportion to their cognitive, emotional, and social complexity. On this gradualist model, there are differences in moral status among sentient nonpersons and not just between them and persons. (DeGrazia 2007, 323)

One need not hold a cognitive sophistication view to endorse what I say about phenomenal richness and value. But perhaps, at least for practical purposes, the phenomenal richness view I outline could be thought of as an ally to cognitive sophistication views.

**3 Implications: Mapping value to moral status**

What does this account indicate regarding the moral status of humans, or of what is probably a broader category, persons? I noted at the beginning of this chapter that many begin with the assumption that humans or persons are at least the paradigm cases of entities with moral status. Arguably, common sense morality maintains a commitment to the thought that humans possess higher moral status than the other animals: a commitment that Kagan (2016) and Setiya (2018), among others, have recently defended.

But if, as seems plausible, an entity’s level of moral status co-varies with the phenomenal richness available to them, it is not obvious that humans have higher moral status than many other animals. If humans do have higher moral status, then this will be on the basis of arguments that take seriously mature scientific theories of not only human psychological capacity and architecture, and of the place of consciousness in human mental life, but also of non-human psychological capacity and architecture, and the place of consciousness in non-human mental life. This work is being done – now more than ever (e.g., DeGrazia 2012, Varner 2012, Rowlands 2019, Barron and Klein 2016). Without seeing where the work leads, claims of human specialness seem unjustified.

What does this account indicate regarding the notion of Full Moral Status? If one thinks that the value of phenomenal consciousness is the chief ground of moral status, then there is an argument against the idea of Full Moral Status (for a similar argument, see DeGrazia 2008). It is that the grounds come in degrees, and do not obviously have a stopping point. So if we do grant Full Moral Status to some being, we confront epistemological worries regarding the justification for drawing the line at any particular place. And we also must allow that the capacity for value in a mental life can expand beyond what is required for Full Moral Status. This does not settle matters, of course, but I think it creates a presumption against the need for a notion of Full Moral Status. We might accomplish just as much with a more practical notion of full protection for beings with certain levels of richness in phenomenal consciousness.

But these notions of degrees and levels of moral status raise troubles of their own. What sense can be made of levels of phenomenal richness?

Ignorance – my own in particular and ours in general – precludes any answer I might offer. It seems to me that, while the account of phenomenal richness as I have outlined it does plausibly predict non-linear jumps between system-types – between animals with very different evolutionary histories and psychological capacities, for example – saying how and why would require a fuller theory of the capacities that enhance phenomenal value, and the ways that they do. And much of this work, both empirical and philosophical, remains undone.

It is important, however, to see that the work can be done. Thinking of phenomenal richness in terms of size, complexity, and coherence allows us to at least think about what kinds of psychological and neurological mechanisms might undergird these features. So, there is at least a hope of mapping a subject’s psychological structure to a subject’s capacities for phenomenal value and disvalue.

To get an idea for how this mapping might go, consider work in neuroscience on ‘levels of consciousness.’ Bayne et al. (2016) make a convincing case that this notion papers over significant complexity. They recommend instead a move towards thinking of the global conscious state of a subject at any particular time as being located somewhere in a multi-dimensional state space. A subject’s global conscious state cycles throughout the day into various places in this state space, given what the subject is doing at a time, what a subject is capable of, and various other factors like whether the subject is awake, and their level of arousal.

Understanding the shape and limits of the multi-dimensional state space for some conscious subject, or even for some species of conscious subjects, is difficult. Bayne et al. note a range of outstanding questions for the neuroscience of consciousness. What capacities and mechanisms structure a conscious subject’s global state at a time? What is the appropriate parcellation of the relevant mechanisms? How might some of these mechanisms interact? What explains the variation in a subject’s global conscious state over windows of time?

These are all good questions, and one can expect progress on answers in the coming years. But note that answers to these questions that apply exclusively to humans will be limited in helping us think through issues of moral status. Additional attention must be given to the ways that very different animals come to occupy regions in a broader multi-dimensional state space – a space that affords comparisons between system-types that have different architectures. Some animals may have regions of state space available to them that are unavailable to humans, and this may be relevant to how we think of phenomenal richness. Other animals may occupy regions of space that, for purposes of phenomenal richness, are relevantly similar to the spaces that humans occupy.

Imagining a program for mapping levels of phenomenal richness may engender skepticism in readers. This is a difficult program, and not well mapped. Epistemic problems leer at every choice point. The epistemic difficulties associated with determining whether some entity is conscious, and if so, what level(s) of phenomenal richness might be available to them, run deep (Shepherd 2018b). We have, at present, no consensus regarding accounts (or theories) of phenomenal consciousness in human brains. We have even less regarding the presence and structure of consciousness in animal minds (see Murray forthcoming). Talk of value present in consciousness complicates matters further. So too does talk of levels.

Significant commitment to the importance of the kinds of questions I have outlined, and significant collaboration on answering these questions, will be required if we are to hope for progress. Significant resources will also be required. Leave potential theoretical difficulties aside for a moment. Isn’t the program doomed for practical reasons?

In the next section I consider ways to make a program for mapping levels of phenomenal richness more practical. It helps that theorists are already thinking in the right direction, even if the vocabulary undergirding their talk of making moral status more practical differs in ways that may signal underlying theoretical disagreements. But it may be possible to work towards something like overlapping consensus.

**4 Making phenomenal value practical**

The current state of the science of consciousness is such that the practical decisions we must make and the policies we must formulate and implement regarding the use and treatment of entities in whom consciousness may be present are associated with uncertainty.

One may respond in different ways to this uncertainty. Aversion to the risk of harm seems plausible. So, often, use of a precautionary principle is suggested. Jonathan Birch explains the basic idea.

In broad terms, the idea is clearly that we should not require absolute certainty that a species is sentient before affording it a degree of legal protection. Absolute certainty will never be attained (indeed, the “problem of other minds” suggests it cannot even be attained with respect to human minds), and its absence is not a good reason to deny basic legal protections to potentially sentient animals. (2017a, 2)

In order to guide practical decision making, a precautionary principle must be articulated with some degree of specificity. Birch follows Stephen John (2011) in counseling interpretation of a precautionary principle in terms of two rules. An epistemic rule specifies where we should set our evidential bar for some fact (e.g., that members of some species S are typically phenomenally conscious) connected to some outcome (e.g., that use of members of S for purposes of research would cause significant harm to those members). A decision rule states that when the evidential bar is cleared, the move to action is urgent, or imperative. As Birch puts it, ‘The implication is that the goal of preventing the seriously bad outcome deserves sufficient priority that, once the evidential bar is cleared, it is inappropriate to delay action further while we attempt to weigh the expected costs and benefits of this goal in comparison to other policy goals’ (2017a, 4).

Birch initially proposed that the evidential bar should apply to animals within an order, and that the bar be statistically significant evidence of ‘at least one credible indicator of sentience in at least one species of that order’ (6). But there is clearly room for debate regarding elements of the evidential bar, as commentary on Birch’s article, and Birch’s (2017b) response to this commentary, indicate.

We can hope that work like Birch’s will galvanize attention to these issues, and focus discussion of ways to move forward. Certainly commitment to a practically implementable precautionary principle regarding potentially conscious entities would require significant collaboration between scientists and policy-makers, and would represent significant progress. In order to implement such a principle, we would need to formulate, keep, and update lists of markers of sentience for different animal-types. And we would need a mechanism for moving between growth in knowledge regarding these markers to the formulating and updating of policies implementing this knowledge by various regulatory and legislative bodies.

But this is, in my view, not all. If there is a contribution the phenomenal richness framework can make here, it is in directing attention to two further issues relevant to any precautionary principle.

*4.1 Proportionality*

Colin Klein (2017) points out that we are often forced to make trade-offs when comparing the impact of a policy on different kinds of animal. For example:

There is a chance that decapods are sentient. The [Precautionary Principle] says: avoid using them in research. Yet perhaps decapod research could help cure cancer, and thereby prevent untold harm to sentient humans . . . (1)

If the scenario is harm to decapods and benefit to humans, how are we to decide? Klein suggests that the Precautionary Principle ‘requires some reformulation in terms of proportionality’ (Klein 2017, 1). How might this go?

One might envision a regime that takes seriously evidence that illuminates the capacity for different levels of phenomenal richness in different types of animal. This might take the form of general measures of psychological sophistication. Consider the Perturbational Complexity Index developed by Marcello Massimini and many others (see Casali et al. 2013, Massimini and Tononi 2018, Comolatti et al. 2019). This is a measure of the informational sophistication of a thalamocortical response to perturbation by, e.g., transcranial magnetic stimulation. Roughly, the PCI indicates the degree of the thalamocortical system’s causal and functional interrelatedness.

This measure has proven predictively fecund. The PCI is able, for example, to distinguish between wakefulness and non-rapid eye movement sleep in healthy adults. And it can discriminate above chance between differences of humans with traumatic brain injury – between those with unresponsive wakefulness syndrome, those in minimally conscious state, and those with locked-in syndrome, for example.

So we have a neurophysiological measure for consciousness that is sensitive to gradations in a subject’s global conscious state. This is not a measure that could yet index more fine-grained differences between system-types. But it is in-principle evidence that biomarkers of and tests for the global conscious state are possible. We should not rule out the possibility of tests for and markers of, not just sentience in various animals, but also levels of psychological sophistication. Given dependable correlations between psychological sophistication and phenomenal richness in humans – if these be established – we may be able to infer something similar in non-humans.

This raises the possibility of the formulation of an evidential bar that has a chance to display sensitivity to gradations in phenomenal richness. For example, following Birch’s language:

For the purposes of formulating animal protection legislation, there is sufficient evidence that animals of a particular [group] possess phenomenal richness at level X if there is statistically significant evidence, obtained by experiments that meet normal scientific standards, of the presence of at least [some number of] credible indicators of phenomenal richness at level X.

This is rough and painfully speculative. But if proportionality matters morally, then there is reason to work towards a more concrete set of measures that allow us to track levels of sophistication across different types of conscious mental life.

*4.2 The source of phenomenal value*

Much discussion of the presence of consciousness in widely different entities focuses on the presence of any consciousness at all. Sometimes there is an added concern for the capacity to feel pain. This makes sense. It is natural to assume that phenomenal consciousness develops all of a piece, and thus that if consciousness is present in an entity, the parts of consciousness that are morally significant will be present as well. The working assumption is that the structure of phenomenal consciousness in different entities is roughly similar, and that the only differences between entities and animal-types are differences of sophistication.

It may turn out, however, that phenomenal consciousness is not produced by a unified mechanism (cf. Phillips 2018). It may instead be the case that there are multiple ways phenomenal consciousness arises in different systems. And there may be implications for the ethics of consciousness.

Peter Godfrey-Smith has recently discussed the possibility that complex perceptual and complex evaluational capacities could evolve separately in some species.

If complex perception and evaluation are separable, this raises the possibility that there are two kinds of phenomena that we vaguely group as ‘subjective experience,’ both of which are present in us but which are distinct in principle and sometimes found separately. (Godfrey-Smith 2019, 14)

Godfrey-Smith suggests that there is some evidence that spiders have sophisticated perceptual capacities, but that they ‘‘score low’ with respect to evidence for motivating feelings’ (14). Godfrey-Smith comments: ‘Perhaps they are sophisticated trackers of the world but motivationally robotic’ (14). Godfrey-Smith further suggests that the opposite may be the case for gastropods. If one accepts something like the phenomenal richness framework I have outlined, one might think that gastropods should be prioritized over spiders in whatever protective legislation might apply.

But the phenomenal richness framework is not mandatory. One might, in the end, favor a neutral view that seeks to protect animals on the basis of sentience alone. Or one might reject the overarching concern with phenomenal consciousness in favor of functional marks of sophistication alone (see Levy 2014).

Whatever view one takes, important practical implications follow. It has been my aim here to suggest that a precautionary principle should be sensitive, not only to the presence of phenomenal consciousness, but to the parts of phenomenal consciousness that are of primary value.

**5 Conclusion**

The chief themes of this discussion are as follows.

First, we need a theory of the grounds of moral status that could guide practical considerations regarding how to treat the wide range of potentially conscious entities with which we are acquainted – injured humans, cerebral organoids, chimeras, artificially intelligent machines, and non-human animals. I offer an account of phenomenal value that focuses on the structure and sophistication of phenomenally conscious states at a time and over time in the mental lives of conscious subjects.

Second, we need to map a theory of moral status onto practical considerations. I prefer the precautionary framework proposed by many, and fruitfully precisified recently by Birch. I have suggested that in addition to further discussion surrounding the evidential bar for attributing consciousness to different types of entities, more discussion is needed regarding how value and moral status may vary across different entity-types, and regarding the sources of value in an entity’s mental life.

**References**

Barron, A. B., and Klein, C. (2016). What insects can tell us about the origins of consciousness. *Proceedings of the National Academy of Sciences* 113(18): 4900-4908.

Bayne, Tim (2010). *The Unity of Consciousness*. Oxford University Press.

Bayne, T., Hohwy, J., and Owen, A. M. (2016). Are there levels of consciousness? *Trends in cognitive sciences*, *20*(6), 405-413.

Bentham, Jeremy (1996). *An Introduction to the Principles of Morals and Legislation: The Collected Works of Jeremy Bentham*. Oxford University Press.

Birch, J. (2017a) Animal sentience and the precautionary principle. *Animal Sentience* 16(1).

Birch, J. (2017b). Refining the precautionary principle. *Animal Sentience* 16(20).

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., ... and Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science translational medicine*, *5*(198), 198ra105.

Comolatti, R., Pigorini, A., Casarotto, S., Fecchio, M., Faria, G., Sarasso, S., Rosanova, M., Gosseries, O., Boly, M., Bodart, O., Ledoux, D., Brichant, J., Nobili, L., Laureys, S., Tononi, G., Massimini, M., G. Casali, A.G. (2019). A fast and general method to empirically estimate the complexity of brain responses to transcranial and intracranial stimulations. *Brain Stimulation* 12(5): 1280-1289.

DeGrazia, D. (2007). Human-Animal Chimeras: Human dignity, moral status, and species prejudice. *Metaphilosophy*, *38*(2‐3), 309-329.

DeGrazia, David (2008). Moral status as a matter of degree? *Southern Journal of Philosophy* 46 (2):181-198.

DeGrazia, David (2012). *Taking Animals Seriously: Mental Life and Moral Status*. Cambridge University Press.

DiSilvestro, Russell (2010). Human Capacities and Moral Status. *Bioethics* 108:165-199.

Glover, J. (2006). The sanctity of life. In H. Kuhse & P. Singer (Eds.), *Bioethics: An anthology* (pp. 266–275). Hoboken: Blackwell.

Godfrey-Smith, P. (2019). Evolving across the explanatory gap. *Philosophy, Theory, and Practice in Biology* 11(1): 1-24.

Hill, Christopher S. (2014). Tim Bayne on the Unity of Consciousness. *Analysis* 74 (3):499-509.

Jaworska, A. and Tannenbaum, J. (2014). Person-Rearing Relationships as a Key to Higher Moral Status. *Ethics* 124 (2):242-271.

John, S. (2011). Risk and precaution. In Dawson, A. (Ed.), *Public Health Ethics: Key Concepts and Issues in Policy and Practice*, pp. 67-84.

Kagan, S. (2016). What's Wrong with Speciesism? *Journal of Applied Philosophy* 33 (1):1-21.

Kahane, G. and Savulescu, J. (2009). Brain damage and the moral significance of consciousness. *Journal of Medicine and Philosophy*, *34*(1), 6-26.

Klein, C. (2017) Precaution, proportionality and proper commitments. *Animal Sentience* 16(9).

Korsgaard, C.M. (2013). Kantian Ethics, Animals, and the Law. *Oxford Journal of Legal Studies* 33 (4):629-648.

Kriegel, U. (2019). The Value of Consciousness. *Analysis*, *79*(3), 503-520.

Lee, A.Y. (2019). Is consciousness intrinsically valuable?. *Philosophical Studies*, *176*(3), 655-671.

Levy, N. (2014). The Value of Consciousness. *Journal of Consciousness Studies* 21 (1-2):127-138.

Liao, S.M. (2010). The Basis of Human Moral Status. *Journal of Moral Philosophy* 7 (2):159-179.

McMahan, J. (2002). *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.

Massimini, M. and Tononi, G. (2018). *Sizing up Consciousness: Towards an objective measure of the capacity for experience*. Oxford University Press.

Murray, S. (forthcoming). A case for conservatism about animal consciousness. *Journal of Consciousness Studies*.

Phillips, I. (2018). The methodological puzzle of phenomenal consciousness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1755), 20170347.

Rowlands, M. (2019). *Can Animals Be Persons?*. Oxford University Press.

Sebo, Jeff (2017). Agency and Moral Status. *Journal of Moral Philosophy* 14 (1):1-22.

Setiya, K. (2018). Humanism. *Journal of the American Philosophical Association*, *4*(4), 452-470.

Shepherd, J. (2018a). *Consciousness and Moral Status*. Routledge.

Shepherd, J. (2018b). Ethical (and epistemological) issues regarding consciousness in cerebral organoids. *Journal of medical ethics*, *44*(9), 611-612.

Siewert, Charles (1998). *The Significance of Consciousness*. Princeton University Press.

Singer, P. 2011. *Practical Ethics, 3rd Edition*. Cambridge University Press.

Tooley, M. (1972). Abortion and infanticide. *Philosophy and Public Affairs* 2 (1):37-65.

Varner, G.E. (2012). *Personhood, Ethics, and Animal Cognition: Situating Animals in Hare's Two Level Utilitarianism*. Oxford University Press.

1. Research on this chapter was supported by European Research Council Starting Grant 757698, under the Horizon 2020 Research and Innovation Programme, and contributes to the project Rethinking Conscious Agency. The author would also like to thank the CIFAR Azrieli program on Mind, Brain, and Consciousness for support. [↑](#footnote-ref-1)
2. Although we differ on specifics, this seems to be the approach of, for example, Peter Singer (2011), as well as David DeGrazia (2007, 2008). [↑](#footnote-ref-2)