

Animal Rights and Conscious Experience

Uriah Kriegel

Arizona

Abstract/Introduction

Perhaps the most influential case for animal rights (see, most notably, Singer 1975) is consequentialist in spirit and relies on two premises. The first is philosophical: the right moral action in every circumstance is that which maximizes pleasure and minimizes pain. The second premise is empirical: that many animals – certainly most animals we have dealings with in our everyday life, including animals used in the food and clothing industries – can experience pleasure and pain. Discussions of this case for animal rights, however, tend to lag behind work on the nature of conscious experience, including experience of pain and pleasure, in the philosophy of mind and cognitive science. Typically, the understanding of conscious experience in those discussions does not go much beyond the folk, pre-scientific understanding. Fully a generation of research in the philosophy and science of consciousness remains unacknowledged. The main purpose of the present paper is to consider the implications for animal rights a more current understanding of conscious experience might have. In particular, I will argue in the bulk of the paper (§§1-3) that on the most plausible account of conscious experience, far fewer animals can experience pleasure and pain than consequentialist proponents of animal rights typically suppose. On the other hand, towards the end of the paper (§4) I will describe an alternative, non-consequentialist ethical framework (combining Kantian and virtue-ethical elements) and argue that it grants animals capable of conscious experience more expansive rights than consequentialist proponents of animal rights

typically grant. The cornerstone of this non-consequentialist framework is the thought that the virtuous agent is s/he who has the stable and dominating disposition to treat all conscious animals, including non-human conscious animals, as ends and not mere means. The upshot is a picture where fewer animals enjoy more expansive rights.

1. Sentience, Consciousness, and Inner Awareness

The consequentialist case for animal rights I have described at the opening is couched in terms of pain and pleasure. One often finds the case put in those terms, but sometimes it is also put in terms of suffering and joy (or enjoyment). There are two ways to understand the relationship between pain and suffering, pleasure and joy. One construes suffering as a deeper, more involved, perhaps more intellectual kind of pain, and correlatively joy as a deeper and more cognitive form of pleasure. The other focuses not on depth but length: suffering is a prolonged, systematic, and stable exposure to pain, joy a prolonged and stable kind of pleasure. In any case, it appears that suffering and joy can be *analyzed* in one way or another in terms of pain and pleasure. For this reason, I will focus henceforth on the latter.

Let us start by distinguishing two notions of pleasure and pain. David Chalmers (1996 Ch.1) argues that mental terms typically lead a 'double life' – they have a *psychological* life and a *phenomenological* life: they can be taken to denote either states conceived of in terms of their psychological role or states conceived of in terms of their phenomenal character. To conceive of a mental state in terms of its psychological role is to conceive of it as essentially the kind of state that has a certain functional role within the subject's overall psychology, that is, the kind of state that has certain typical causes and effects. These causes and effects can be restricted to sensory inputs, behavioral outputs, and other mental states, but can also be taken more widely to include distal causes and effects in the subject's

environment. To conceive of a mental state in terms of its phenomenal character is to conceive of it as essentially the kind of state that has a certain felt quality, the kind of state there is something it is like for the subject to be in.

At the background is a general distinction between two conceptions of mind. The psychological conception of mind characterizes mental phenomena third-personally in terms of their causal relations to each other and to the environment; the phenomenological conception characterizes them first-personally in terms of their phenomenal or subjective feel. The former focuses on the *mechanical* dimension of mental life, the latter on its *experiential* dimension. Both dimensions are real and central to the nature of mind, so there is no question here of which conception is 'more accurate' – both are important.¹

Consider now the implication for pleasure and pain. Mental states can be classified as pain, say, either because (roughly) they are typically caused by harmful stimulation (e.g., tissue damage) and cause aversive reaction (e.g., cringing), or because they feel that particular unpleasant way – they hurt. Correspondingly, the term 'pain' can be used to denote either any mental state that typically is suitably caused by harmful stimulation and causative of aversive reaction or any state that feels that particular unpleasant way. These classification criteria not only are distinct, but also do not coextend. Suppose a subject experiences a toothache, but upon noticing a fire in the kitchen, starts attending to the fire and is thus too distracted to experience the toothache, which unfortunately reappears as soon as the kitchen drama is over. Should we say that this subject had a single continuous toothache, which went unconscious at some point and was then consciously experienced again, or that the subject had two distinct pains, separated by a painless interval? The answer, of course, is that we should say both, but in different senses of 'pain.' In the *psychological* sense, there is only one pain. This is because there is unconscious pain – a mental state which is not experienced by its subject but which nonetheless plays the pain role in the subject's psychology.² In the *phenomenological* sense, by contrast, there are two pains, and there is no such thing as non-experienced pain – as soon as the early pain is phenomenally extinguished, it goes

out of existence, and a new one reemerges when the phenomenology of pain reappears later.³

The upshot is that we need a distinction between *psychological pain*, which is pain conceptualized through its mechanical or functional profile, and *phenomenological pain*, which is pain conceptualized through its experiential or phenomenal profile. This distinction has obvious implications for the consequentialist thesis that our decision making should attend exclusively to the pains and pleasures our actions are likely to bring about.⁴ Let us state the relevant consequentialist thesis as follows:

(CT) In deciding what actions to perform, one ought always to attempt to maximize the pleasures and minimize the pains consequent upon one's actions.

For the purposes of a consequentialist defense of animal rights a weaker thesis – taking pain and pleasure to be sufficient conditions for moral relevance, say – may be enough. But (CT) will do for the sake of illustration. The central point I want to make here is that the effect of the above distinction between two notions of pain and pleasure is to render CT ambiguous. For CT now admits of three different interpretations, calling on us to minimize/maximize either (i) psychological pain and pleasure, (ii) phenomenological pain and pleasure, or (iii) both psychological and phenomenological pain and pleasure.

It seems clear that the second interpretation is greatly preferable to the first. Although the issue is not trivial, it is fairly obvious that, other things being equal, producing phenomenological pleasure in someone is commendable while producing phenomenological pain is criticizable. (The issue is not trivial because it is not obvious *why* that is the case; but it does seem obvious *that* it is the case.) By contrast, it is hard to see how producing psychological pleasure and pain are

supposed to be commendable or criticizable. Suppose that, perhaps *per impossibile*, a race of zombies was discovered (and ascertained to be zombies) who were functionally indistinguishable from us, in the sense that the functional organization of their psychology would be identical to ours. These zombies would have internal states with the same functional role as our pleasures and pains, but would never consciously experience pleasure or pain. In other words, they would undergo psychological pleasures and pains but not phenomenological pleasures and pains. On the face of it, there is no reason to think that causing these zombies the internal states corresponding to our pleasure would be worth the effort and investment, or that causing them the internal states corresponding to our pain experiences would be worth the effort of avoiding. If the zombies do not *experience* their pleasures and pains, but simply *host*, so to speak, internal states with the pleasure/pain functional role, their internal states do not have any moral weight.

The point can be brought out in a different way. Consider that it is only when a pain is experienced that it hurts the person undergoing it – that it is only when a pain is experienced that the person is *pained by* the pain; and likewise, that only when a pleasure is experienced that it *pleases* its subject. If there are internal states that do not pain or please the persons undergoing them, but which deserve the names ‘pain’ and ‘pleasure’ for some other reason (e.g., due to their functional role), those states do not carry any moral weight. For it is the pain and pleasures *of persons*, or at least *of subjects*, that interest us.⁵ After all, if there were self-standing pains and pleasures floating about, as it were, without belonging to anyone, we would not find them to be appropriate targets of moral concern.

Notice now that since we found psychological pains and pleasures to be morally insignificant, it would seem that the second interpretation of CT is preferable not only over the first one, but also over the third one, that which recommends maximizing both phenomenological and psychological pleasure and minimizing both phenomenological and psychological pain. That is, the moral weight of CT is better captured through the following principle:

(CT+) In deciding what actions to perform, one ought always to attempt to maximize the phenomenological pleasures and minimize the phenomenological pains consequent upon one's actions.

I take this point to be in line with the standard consequentialist defense of animal rights, as can be seen by its common appeal to the notion of *sentience* (Singer 1975).⁶ The term 'sentience' is hard to associate with merely psychological or functional states; rather, it intimates a phenomenal, experienced state. Animals may indeed have rights insofar as they are sentient, but sentience is probably associated only with the capacity to be in internal states with a phenomenal character, an experiential dimension – not with the capacity to be in internal states with a functional role (this latter capacity being present even in carburetors, which are of course insentient).

This is important because there is a central debate in the philosophy of mind about one specific question that arises with phenomenal states but not psychological ones. The question is whether phenomenal states are necessarily states of which their subjects are aware. Answering in the positive, Rosenthal (1993) formulates what he calls the *transitivity principle*: conscious states are states we are aware of.⁷ The relevant sense of 'conscious' here is the phenomenological sense, so the claim is that it is necessary for a mental state's being phenomenally conscious that its subject be aware of its occurrence. After expanding somewhat on the transitivity principle, I will argue that its truth is likely to have significant implications for the scope of animal sentience.

The transitivity principle is appealing to philosophers who adopt higher-order or self-representational theories of phenomenal consciousness. According to the former, a mental state is conscious only if the subject entertains a higher-order representation of it. The higher-order representation may be either a thought

(Rosenthal 1990, 2002) or a quasi-perceptual state (Armstrong 1968, Lycan 1996), but either way it is necessary for its target's being experienced. According to self-representational theories, a mental state is phenomenally conscious only if it represents itself (Brentano 1874, Caston 2002, Kriegel 2003, 2009, Kriegel and Williford 2006, Williford 2006). On both views, a phenomenally conscious state must be represented. What motivates this is the transitivity principle: the idea that a mental state cannot qualify as conscious if its subject is utterly unaware of it. Mental states we are unaware of – mental states we have no idea are taking place in us – are unconscious. The awareness of our phenomenal states need not be reflective and explicit, but some kind of awareness must exist.⁸

The transitivity principle is quite intuitive, but some philosophers are willing to reject it. In particular, proponents of so-called first-order representational theories, as well as functionalist theories, are committed to its falsity. According to the former, a mental state is phenomenally conscious if it represents the right environmental features, or represents them in the right way (Dretske 1995, Tye 2000). According to the latter, a mental state is phenomenally conscious if it has the right functional role, perhaps one that guarantees its global accessibility to consuming modules (Baars 1988) or verbal report (Dennett 1991).⁹ Since a mental state can represent any environmental features and play any functional role without its subjects being aware of it, proponents of such theories must reject the transitivity principle.

The debate over the merits of the transitivity principle is too complex to broach here. I have argued for it in some detail elsewhere (Kriegel 2009), but replicating that argumentation here would take us too far afield. There are only two points I want to make here.

The first is to point out a conditional: *if* the transitivity principle is true, then it is likely that far fewer animals are sentient in the relevant sense than Singer and other proponents of animal rights tend to claim. The relevant sense of sentience, recall, is that in which a creature is capable of experiencing phenomenological pain

and pleasure. If the capacity to undergo such experiences requires the further (and much more cognitively sophisticated) capacity to be aware of one's pain and pleasure, if only non-reflectively, then given that relatively few animals have that further capacity, we can expect relatively few animals to have the capacity to experience phenomenological pain and pleasure. Which animals we can reasonably expect to have that capacity will be discussed in the next section.

The second point I would like to make is that even if we accept that there are phenomenally conscious states of which the subject is unaware, it is not clear that such states would require our moral attention in the way phenomenal states we are aware of do – for reasons similar to those brought up against the moral weight of psychological pain and pleasure. Thus, suppose that, as some opponents of the transitivity principle have suggested, there could be sub-personal mental states, buried deep in subjectively inaccessible modules, that were nonetheless phenomenally conscious (Block 2007). It is not at all clear why we should be concerned not to inflict such states on others, or even be concerned that we ourselves not undergo such states. If there could be a phenomenological pain of which I were completely unaware, why should I try to avoid it? After all, what we are asked to imagine here is a situation in which I undergo a pain state but am not *pained by it*. Given this specification, and stipulating that this pain state by which I am unpained lacks any further adverse consequences (which might pain me), it would seem irrational for me to fear it or invest in avoiding it. Indeed, other things being equal it would be rational for me to accept great pain of which I am utterly unaware, and which thus does not pain me, in return for the slightest pleasure of which I *am* aware, and which therefore pleases me.

To repeat, as mentioned above, my own position is that the transitivity principle is true, and so phenomenally conscious pains and pleasures must be ones the subject is aware of, and is thus pained and pleased by. Against this background, the supposition that they can occur without paining and pleasing their subjects is nonsensical. But my present point is that even if we indulge the objector's supposition that such phenomenal pains and pleasures make perfect sense, and

there were phenomenal pains and pleasures of which subjects were completely unaware, such pains and pleasures would not carry the sort of moral significance we the folk (as well as committed consequentialists) typically assign to them. In other words, if the objector were right, we would have to revise the consequentialist principle as follows:

(CT2) In deciding what actions to perform, one ought always to attempt to maximize the phenomenological pleasures of which their subject is aware and minimize the phenomenological pains of which their subject is aware consequent upon one's actions.

(CT+) would still be a consequentialist principle, of course, but it is (CT2) that would be the *plausible* consequentialist principle.

Note, in any case, that it is clearly implausible to impose a transitivity requirement on pain and pleasure understood *psychologically* rather than *phenomenologically*. A mental state can have a role in the subject's psychology without the subject being aware of it. Consider psychological pain: a mental state can certainly be caused by harmful stimulation and cause aversive reaction without the subject being in any way aware of its occurrence. This means that in one sense of the term 'pain,' it is indeed plausible that pain can occur without awareness of that pain, and moreover, that many – for all we know, all! – animals undergo pain in that sense. This may create the temptation to say that most or all animals are sentient. But of course that temptation, so grounded, trades on an ambiguity in the terms 'pain' and 'pleasure.'¹⁰ Pain and pleasure can occur without their subject's awareness of them, but plausibly, this is the case only in the psychological sense of the terms, and the psychological sense is not the one relevant to sentience. What is relevant to sentience is the phenomenological sense of the terms 'pain' and 'pleasure,' but in that sense, pain and pleasure do require their subject's awareness

of them, or if they do not, their occurrence without the subject's awareness does not carry the relevant moral significance.

All this matters, as noted above, because it is an open question how many animals are capable of being aware of their own internal states, in particular their phenomenal pains and pleasures. Let us call the kind of awareness one has of one's phenomenally conscious states *inner awareness*. The question before us is how many animals are capable of inner awareness – and, as a corollary, how can we know which are. I turn to discuss these matters in the next section.

2. The Question of Animal Consciousness

From now on, I will assume that the transitivity principle is true, and so that phenomenal consciousness requires inner awareness. Given that sentience is best understood as phenomenal consciousness, it follows that sentience cannot occur without inner awareness. The question of the scope of animal sentience is thus constrained by the question of the scope of animal inner awareness: there can be no more sentient animals than animals capable of inner awareness. The question is which animals are capable of inner awareness of their internal states. And since such inner awareness requires the capacity to form representations of those internal states, the question is which animals are capable of forming representations not only of their environment, but also of their own internal states.

Positions on this issue vary quite a bit. On the one extreme is Peter Carruthers' (1999, 2004) view that no non-human animal has that capacity, and therefore that no non-human animal is phenomenally conscious. Carruthers reaches this conclusion by taking a side on a scientific debate on the extent of so-called metacognition. There is an ongoing debate among psychologists concerning the extent to which animals are capable of assigning to themselves (or for that matter, to conspecifics) false beliefs and other misrepresentations of the external world.¹¹

The main protagonists in this debate are Michael Tomasello, who argues that at least apes are capable of such ascription, and Daniel Povinelli, who argues that they are not. Carruthers takes Povinelli's side, and thus adopts the view that even apes have no metacognitive capacities (see Tomasello and Call 2005, Povinelli et al. 1994). This suggests that they cannot form representations of their own internal states. Since he also subscribes to a variety of higher-order theory of consciousness (Carruthers 2000), he concludes that even apes are unconscious and insentient. The view is extremely counter-intuitive, he admits, but (a) it is forced on us by the relevant scientific findings (coupled with certain philosophical commitments, of course) and (b) our strong contrary intuitions can be explained as an artifact of projection (Carruthers 2004).

There is much that is questionable in this reasoning, as we will see momentarily.¹² At the same time, even if apes do show signs of metacognition, it is clear that evidence of metacognitive capacity in non-primates is extremely sparse. Correlatively, the capacity for mirror self-recognition, which is taken by some to be the best operational definition of self-awareness (Gallup 1975) has been demonstrated and replicated in only four non-human animals: the chimpanzee, the orangutan, the bottlenose dolphin, and the Asian elephant.¹³ Even allowing that mirror self-recognition is too demanding a criterion for self-awareness, and false belief ascription too demanding a criterion for inner awareness, there is no doubt that these experimental findings, combined with the thought that sentience requires inner awareness, challenge the pre-theoretic intuition that sentience is widespread in nature.

Consider the case of our favorite pets, cats and dogs. We certainly assume in everyday life that they are sentient creatures, capable of feeling phenomenal pain and pleasure. But even if we do not take their inability to ascribe false beliefs to themselves (or recognize themselves in the mirror) to provide evidence for their lack of inner awareness, and therefore (given the transitivity principle) lack of sentience, this still leaves us without any positive evidence for the *presence* of inner awareness and sentience. Moreover, it puts a cloud over their prospects: if sentience

requires phenomenal consciousness, phenomenal consciousness requires inner awareness, inner awareness requires metacognition (or more generally representation of one's own internal states), and none of the standard tests for metacognition turns up evidence for the presence of metacognition in our cats and dogs, then we should at least take seriously the possibility that our cats and dogs are not in fact sentient, and do not experience pain and pleasure.

As a matter of psychological fact, it seems virtually impossible to look into one's pet's eyes and seriously consider the possibility that s/he is unconscious and insentient. When I make eye contact with (my favorite live dog) Hailey and detect anticipation in her as I reach for her leash, it is inconceivable to me that I am interacting with a sophisticated automaton. But why is that? Whence our strong intuition that our pets are conscious? Recent studies of the folk's intuitive ascription of consciousness have found that such ascription is based on factors such as *having eyes, being self-moving*, and a number of similar factors (Arico et al. forthcoming). I trust the reader can appreciate that these are not 'deep' indicators of sentience. This casts our intuitions about animal consciousness as sub-optimally reliable. It may well be that the intuition that cats and dogs are conscious is in fact correct, and indeed I will eventually argue that it is. But it is important to appreciate that the reason we have the intuition is *not* the reason that makes probable the sentience of cats and dogs. At most, then, the intuition is *accidentally correct*.

There may be a sounder basis for the ascription of consciousness to our pets than the one that seems to underlie our intuitive ascription. For example, there seems to be ample behavioral evidence: when we accidentally step on our dog's paw, s/he jolts and shrieks in a manner very much indicative of experienced pain. That is, we observe the correspondence between harmful stimulation and aversive reaction in the dog that we have observed before in humans, and conclude that just as when this happens in humans, they are in pain, so the dog must be in pain. But here we should keep in mind that the reference to harmful stimulation and aversive reaction is constitutive of pain only in the *psychological* sense, not in the *phenomenological* sense. So what the dog's behavior provides direct evidence for is

the dog's capacity to undergo *psychological* pain. It does not provide direct evidence for the dog's capacity to undergo *phenomenological* pain.

Perhaps we can justifiably take this behavior to provide *indirect* evidence for phenomenological pain. The reasoning would be this: since aversive reaction to harmful stimulation is constitutive of psychological pain, the dog's behavior provides direct evidence for the presence of psychological pain in the dog; and since there is certainly a robust correlation between psychological and phenomenological pain in us, we are entitled to infer that they do in other creatures (pending evidence to the contrary), and so we are entitled to infer that the dog's psychological pain correlates with a phenomenological pain. Thus we obtain indirect evidence for the presence of phenomenological pain in the dog from the dog's behavior (see Lurz 1999).

However, there is a possible disanalogy here. Suppose for the sake of argument that humans are capable of inner awareness but dogs are not. Then the fact that there is correlation between psychological and phenomenological pain in humans would not underwrite a strong inference to the existence of such correlation in dogs. For it may well be that what phenomenological correlates not with psychological pain but with inner awareness of psychological pain. Indeed, this is the natural view to adopt if we accept the transitivity principle.¹⁴

For all these reasons, it seems to me that adopting the transitivity principle does put in question the sentience of many animals we are intuitively inclined to consider sentient – including our pets. In the remainder of this section, I consider three type of consideration we might cite in resisting this untoward result, so as to tilt the balance of evidence in animals' favor. In the next section, I will expand on the third, which I take to be the most useful guide in determining the scope of animal consciousness.

The first option is to contest the experimental consensus regarding metacognition, claiming on alternative experimental grounds that metacognition is much more rife than commonly thought. Consider the ascription of false beliefs,

starting with the case of humans. It has been something of a consensus among developmental psychologists for about two decades now that toddlers do not acquire the capacity to ascribe false beliefs to themselves before the age of three (Perner et al. 1987, Gopnik and Astington 1988). However, Fodor (1992) has argued that the false belief tasks used in the central experimental paradigms are simply too difficult for infants *as tasks*, but that toddlers and infants may nonetheless have the relevant capacity. This claim may be borne out in recent work by Baillargeon, who has found (through a series of ingenious experimental designs) compelling evidence of understanding of false belief in infants as young as eight months (Onishi and Baillargeon 2005, Song and Baillargeon 2008). This would certainly suggest that a similar situation could obtain with respect to experimental work on metacognition *in animals*: that tasks used thus far may have been too difficult, and that clever designs may reveal signs of metacognition in much less sophisticated animals than even apes.

The second option is to claim that although the metacognitive capacities tested for in the central studies are relatively sparse, the bare capacity for inner awareness requires much more meager metacognitive capacities. Rosenthal (2000) adopts this strategy. He claims that many animals may be incapable of ascribing themselves false beliefs while nonetheless being capable of ascribing to themselves a mental state, because the concept of a mental state in general is much simpler than the concept of belief more specifically. Researchers have sought the ability to ascribe false beliefs because it is thought that only the ascription of a false belief indicates an appreciation of the difference between appearance and reality, between the way things are and the way one represents them to be. But although *our* ascriptions of beliefs to ourselves involve this appreciation, it may well be, according to Rosenthal, that other animals' do not. There is no reason to take such appreciation to be *constitutive* of inner awareness as such. Thus a metacognitively challenged animal (if you will) may lack the concept of belief but possess the concept of truebelief, and such an animal could still be aware of her internal states, even if not in the way we adult humans are. Furthermore, as I point out in Kriegel

2009 Ch.5, there is no contradiction in the notion of an animal that is aware of its internal states *non-conceptually*, so inner awareness may not require *any* conceptual capacities.

The third option is to insist that even if there is no behavioral evidence of animal inner awareness whatsoever, there may nonetheless be *neural* evidence of it. More generally, we may look for neural evidence of consciousness in non-human animals, and when we find it, take it quite reasonably to override behavioral counter-evidence. I expand on this in the next section.

3. An Empirical Procedure

In this section, I apply, somewhat speculatively, what I take to be the strongest empirical procedure for determining which animals are capable of sentience. The procedure is fairly straightforward: we ought to seek the neural correlates of consciousness/sentience in creatures that are unquestionably conscious, then attempt to establish which other creatures, if any, exhibit neural structures relevantly similar to these. Several clarifications are in order, however, before we can undertake application of this procedure.

First, which animals are ‘unquestionably’ conscious depends on what we take ‘unquestionably’ to mean. I suggest that the best understanding of ‘unquestionably’ is as follows: it is unquestionably the case that p just in case any claim that $\sim p$ would not be taken seriously within the relevant community of inquiry (where this could be operationalized in terms of publications in peer-reviewed journals, presentations in academic gatherings, or some such indicators). In the present case, then, a creature is unquestionably conscious just in case the claim that it is not conscious would not be taken seriously within the neuropsychological community. Since as a matter of fact even Povinelli’s claim that chimpanzees are unconscious has been taken seriously, but no claim that normal adult humans are unconscious has ever

been taken seriously, I suggest that we consider normal adult humans as our paradigmatically unquestionable instance of a conscious creature.¹⁵

Secondly, what makes a brain structure ‘relevantly similar’ to another is also problematic, especially in the context of cross-species comparison. Neuroanatomical comparison, in particular, may be neither reliable nor feasible. Better to focus on comparisons of structures functionally individuated. Thus, one brain *region* may be different in terms of its spatial location from another and yet qualify as a similar brain *structure*, namely, if the two regions perform the same neural function.

Thirdly, the fact that a brain structure is relevantly similar to one that underlies consciousness does not *guarantee* that it too underlies consciousness, and likewise the fact that it is relevantly *dissimilar* to one that underlies consciousness does not guarantee that it does *not* underlie consciousness. However, such relations of relevant similarity and dissimilarity do offer non-deductive evidence of underlying consciousness (or not), by supporting an inference to the best explanation to that effect. Such an inference makes its conclusion *probable*, though it does not guarantee its truth. Thus if brain structure S_1 underlies consciousness in creature C_1 , where C_1 is unquestionably conscious, and structure S_2 is relevantly similar to S_1 , then it is probable that S_2 underlies consciousness in C_2 ; and if S_2 is relevantly dissimilar to all brain structures of all unquestionably conscious creatures, then S_2 probably does not underlie consciousness in C_2 .

Incorporating these clarifications into our understanding of the empirical procedure I described leads to the following methodological principle:

(M1) For any brain structure S of a creature C , if there is a brain structure S^* of normal human adults, such that (i) S^* underlies consciousness in normal human adults and (ii) S is neurofunctionally similar to S^* , then (probably) S underlies consciousness in C ; and if there is no brain

structure S^* that satisfies (i) and (ii), then S probably does not underlie consciousness in C .

We can then consider that an animal is probably unconscious when none of its brain structures is neurofunctionally similar to one that underlies consciousness in an unquestionably conscious creature. That is:

(M2) For any creature C , C is probably conscious if and only if there are brain structures S and S^* , such that (i) S belongs to C , (ii) S^* underlies consciousness in normal human adults, and (iii) S is neurofunctionally similar to S^* .

With (M2) in place, we can generate a list of probable conscious animals as soon as we establish two things: first, which brain structure(s) underlie consciousness in normal human adults, and second, which animals have brain structures neurofunctionally similar to those structures in the human brain. Our current state of knowledge with respect to both questions is radically incomplete, but I end this section with cursory remarks on both.

The scientific debate over the neural correlates of consciousness is very much alive, and no consensus has emerged as yet. However, two main camps have emerged. The minority camp holds that the neural correlates of consciousness are present in low-level parts of sensory cortex (Zeki and Bartels 1999, Gray 2004). The majority camp holds that it involves also higher, more sophisticated brain areas, in particular prefrontal areas. One particularly influential hypothesis designates a frontoparietal network as underlying consciousness (Dehaene et al. 2003). Another adverts to inferotemporal cortex and superior temporal sulcus as underlying at least

visual consciousness (Leopold and Logothetis 1996). Yet another – the most experimentally sophisticated, in my view – homes in on the dorsolateral prefrontal cortex (Lau and Passingham 2006, see also Sahraie et al. 1997). The hypothesis I find most compelling is the *COI hypothesis*, according to which the neural correlate of consciousness in normal adult humans is neural synchronization with activity in dorsolateral prefrontal cortex (Kriegel 2007). This is not the place to compare the evidence for all these hypotheses, or even get clear on what exactly they mean. Let me only quickly explain what the COI hypothesis says.

The dorsolateral prefrontal cortex is known to be involved in several neural functions, but the one relevant to the COI hypothesis is its involvement in higher-order monitoring of lower brain areas' activity (Gehring and Knight 2000). Thus the dorsolateral prefrontal cortex is relevant to the neural correlates of consciousness in virtue of generating higher-order representations. If we take such higher-order representations to implement inner awareness, it would seem to follow that the dorsolateral prefrontal cortex underlies inner awareness. Meanwhile, neural synchronization is widely accepted to implement as a mechanism that binds different representations in the brain (von der Malsburg 1981): although the shape and color of an object are represented in two different areas of the brain, the brain represents that they belong to the same object by synchronizing the neural activity in those two areas. Synchronization of a representation in the brain with dorsolateral prefrontal activity is intended to capture the thought that a representation R becomes conscious when (i) there is a higher-order representation of R and (ii) R is integrated, or unified, with this higher-order representation of it. The result is cross-order integration, or COI for short.

Let us adopt this hypothesis for the sake of discussion. Then given (M2), we should recognize an animal as probably conscious just in case it meets two conditions. The first is that the animal have a brain structure relevantly (neurofunctionally) similar to the human dorsolateral prefrontal cortex. Since the relevant neural function of the dorsolateral prefrontal cortex is its ability to perform higher-order monitoring and produce higher-order representations, this is the

neural function we should look for in non-human animals' brain structures in order to qualify them as 'relevantly similar' to the human dorsolateral prefrontal cortex. The second condition is that the animal have some mechanism for integrating representations, akin to the synchronization mechanism in human.

It is almost certain that most animals capable of any minimal type of cognition utilize a synchronization mechanism (or a functionally analogous mechanism) to integrate disparate bits of information about their environment. Without such capacity they would be hard-pressed to navigate their environment with any success. So most animals would certainly meet the second condition I have described. The more interesting question is which animals are likely to meet the first condition, that is, to have a brain structure involved in higher-order monitoring of brain activity.

Research on higher-order monitoring in animals is virtually non-existent, so no firm claims can be made regarding this question. It is clear, however, that many animals that consequentialist proponents of animal rights claim are sentient have no dorsolateral prefrontal cortex and are unlikely to exhibit higher-order monitoring. Certainly no invertebrates or fish can be reasonably thought to pass this test, as they do not even have cortex, but it is hard to believe that amphibians, reptiles, and birds, and thus any non-mammals, would either.¹⁶ Given (M2), this suggests that all non-mammals are probably phenomenally unconscious, that is, insentient. This means that (CT+) does not apply to them: for they do not experience pleasure or pain.

It is possible, of course, that animals lacking a dorsolateral prefrontal cortex still perform higher-order monitoring: the function is implemented by a different brain area. However, to repeat, although this is perfectly possible, an inference to the best explanation casts it as improbable, since it is improbable that wildly different brain regions implement the same functions in the actual world. (Indeed, this is why (CT+) is stated merely in terms of probability.)

I conclude this section somewhat open-endedly, insofar as which animals are conscious and sentient would depend, if I am right, on as yet non-existent research on higher-order monitoring in animals. If I had to, however, I would predict that only mammals will turn out to exhibit higher-order monitoring, and therefore that only mammals will turn out, against the background of the methodological principle I have adopted here, to be conscious.¹⁷ By (CT+), this requires us to take the pleasure and pain of mammals, but not of non-mammals, into consideration.

4. A Non-Consequentialist Variant

I opened this paper with a construal of the standard consequentialist case for animal rights as involving a philosophical premise and an empirical premise. I have spent most of this paper contesting the empirical premise. But although I cannot argue for this here, I think we should reject the philosophical premise as well. In this section, I describe an alternative ethical framework and point out some of its advantages in granting (conscious) animals greater rights.

The immediate alternative to consequentialism is of course deontology, of which the Kantian variety is the best known. For my part, I find the second formulation of the categorical imperative, the 'humanity formula,' the most compelling: one should always act in such a way that one treats humanity, whether others' or one's own, as an end in itself and not merely as a means to some other end. Perhaps because this formulation of the categorical imperative, and Kant's moral philosophy as a whole, restrict moral worth to the sphere of humanity, Kantian defenses of animal rights are hard to come by. But it should be straightforward for a proponent of animal rights with Kantian predilections to modify Kant's dictum to encompass the sphere of the sentient or conscious. Thus let us call the following the *consciousness formula*:

(CF) One should always act in such a way that one treats conscious creatures as ends in themselves and not merely as means to other ends.

There are complicated questions surrounding whether Kant's own reasons for putting forward the categorical imperative can survive this modification. But regardless of how that turns out, it is clear that (CF) is a coherent position, and can therefore be adopted by a non-consequentialist proponent of animal rights.

The case for (CF) is unlikely to be Kantian, then, but I think it should be clear what kind of case it would be, since it speaks to the combination of deontological and animal-rights sensibilities. These sensibilities are well captured in something like the following bit of reasoning: all conscious creatures have moral worth; one ought to treat creatures with moral worth as ends rather than as mere means; therefore, one ought to treat all conscious creatures as ends rather than as mere means.

The ethical framework I would like to adopt here takes its cue from this Kantian approach but gives it a virtue-ethical twist.¹⁸ The moral principle I propose to pin animal rights on is thus neither (CT+) nor (CF), but rather:

(CF+) One should have the stable, dominating disposition to treat conscious creatures as ends in themselves and not merely as means to other ends.

Call this the *virtue-ethical consciousness formula*. I will not argue for it here – that is the proper target of a much larger project – but I will show a couple of ways in which it grants conscious animals more rights than (CT+). Before doing so, however, let me clarify the main terms employed in (CF+).

Start with the notion of a disposition to treat someone a certain way. The metaphysics of dispositions is a contentious area of research, but the first analysis to come to mind is the so-called simple conditional analysis (Ryle 1949, Quine 1960, Choi 2008): x has the disposition to exhibit manifestation M in circumstances C iff: if C were the case, x would exhibit M . Many philosophers today reject the simple conditional analysis, in favor of either a complicated conditional analysis (Lewis 1997), a non-conditional analysis (Fara 2005), or a primitivist account (Molnar 1999).¹⁹ Different accounts of disposition will result in different understandings of the disposition to treat conscious creatures as ends and not mere means. It is easiest, for purposes of exposition, to illustrate this with the simple conditional analysis, so I will stick with it. According to it, a person has that disposition if the following conditional is true of her: in most circumstances, she does treat conscious creatures as ends and not mere means.

Note that this conditional requires that the disposition manifest itself in *most* circumstances. This is what the qualifier ‘stable and dominating’ is supposed to capture: the greater the variety of types of circumstance in which the disposition is manifested, the stabler the disposition, and the greater the number of times the agent manifests the disposition in each type of circumstance, the more dominating the disposition (in the sense that it dominates more of the agent’s behavior). Clearly, on the whole the more stable and dominating the disposition, the more commendable the agent by the lights of (CF+). But it is an interesting question whether our ideal should be that the disposition be *totally* stable and dominating, that is, manifest itself in *every* circumstance. On the one hand, the *phronimos*, or the ideally virtuous agent, surely treats conscious creatures as ends in themselves *always and everywhere*. At the same time, it appears to be a psychologically contingent fact about sub-optimal moral agents such as us that absolute adherence to moral dicta tends to make us dour and intolerant: we run the danger, in our puritanism, of becoming deeply disappointed with humanity (for failing to live up to the moral standards we live up to) and ultimately adopting a derisive and uncompassionate attitude towards our peers. To avoid these pitfalls, it may be wise

for us to demand of ourselves something less than a perfectly stable and dominating disposition.²⁰

The last clarification we ought to offer is of the phrase ‘ends and not mere means.’ Here too, the existing literature is vast and there is no need (or hope) for us to make an original and plausible contribution to it here. It bears stressing, however, that not treating someone as a *mere* means does not entail not treating them as a means. If A treats B both as a means and as an end, A does not treat B as a mere means, in the sense that A does not treat B merely as a means. This is why asking someone for the time does not *necessarily* involve treating them as mere means, although it may, namely, if the agent does not make sure to treat the patient *also* as an end. This raises the thorny issue of what is involved in treating someone as an end (since it cannot be analyzed in terms of not treating them as a means, given that one can treat a person both as an end and as a means). Presumably, it involves the agent’s keeping in mind, in her dealings with the patient, the right psychological attitude (respect, empathy, love, or what have you) toward the patient. For my part, I am tempted by the view that there is a *sui generis* phenomenology of treating someone as an end, and that the treating of someone as an end is to be analyzed in terms of the presence of this phenomenology. Another tempting view is that there is a special kind of affective state, which may be called with some qualification love, that is involved in this.²¹ Again, different positions on this matter will result in different versions of (CF+).

(CF+), so understood, diverges from (CT+) in its practical prescriptions in a number of ways. Perhaps the most striking of these is the fact that, if a conscious animal could be killed without causing pain to it (or anybody else), it is not clear that (CT+) would prohibit the killing. As a matter of fact, conscious animals raised today in the mainstream food industry typically lead miserable lives of ceaseless suffering. However, it is conceivable that some conscious creatures be brought up in quite pleasant circumstances, but for the purpose of eventual (painless) slaughter. Indeed, it is not implausible that this in fact takes place in some less industrialized parts of the world. The consequentialist defense of animal rights seems to me to

have no resources for prohibiting such a practice. The virtue-ethical Kantian framework I have presented clearly does, since the relevant conscious creatures are treated as mere means, insofar as it is impossible to treat someone also as an end in killing them in order to meet one's own gastronomic desires.²²

Similarly, consider the possibility of raising conscious animals, in a humane and painless way, for the purpose of consuming their products (e.g., their milk). Traditional farming without commercial ambition often took this form. Again, it does not seem that (CT+) has the resources to prohibit such a practice, but (CF+) does.²³ Thus, if combined with the view that all animals are conscious, (CF+) would prescribe not only vegetarianism but veganism. For reasons explained earlier, I think it should only be combined with the claim that all *mammals* are conscious, and so veganism does not follow: we are most certainly justified in consuming honey, for instance, since bee brains most certainly do not engage in higher-order monitoring (and if I am right, the same holds also for the products of fish and chickens, e.g., eggs).²⁴ At the same time, according to (CF+), commercial products involving the products of (conscious) mammals are morally tainted one and all: leather jackets, shoes, and belts, fur coats, ivory sculptures, etc. are all to be avoided.²⁵

The more general point to appreciate here is that (CF+) and (CT+) must diverge, since whether an animal is treated as a mere means and whether it is pained are two different matters. Doubtless they overlap quite often in practice, but they are distinct nonetheless: an animal, even a human animal, can be pained without being treated as a mere means (as when we tell a friend a painful truth because it is the right thing to do), and can be treated as a mere means without being pained (as with the well-treated cows). My claim is that in all these cases, it is treating the animal as a mere means, rather than painning the animal, that tracks moral wrongness as such.

These are different facets of the more general fact that (CF+) prohibits treating conscious animals as a resource for and by humans. This general fact, it seems to me, may well capture the moral sensibility that animates the case for

animal rights. Yet there is nothing in the consequentialist approach to animal rights that speaks to this concern. By contrast, the virtue-ethical Kantian approach I have adopted, in giving all conscious animals intrinsic moral worth, ensures that conscious animals do not derive their worth from the value they have for humans. Within this approach, accepting the use of conscious animals for our own purposes under the condition that they are well-treated and are not suffering is akin to defending humane slave owners on the grounds that they offered their slaves a relatively pleasant life. Even if such slaves worked regular hours and were generously compensated, indeed worked less and were better compensated than they would be as free men and women, the very fact that they would be considered the possession of, and a resource for, their owner is what we would find so morally objectionable. More generally, the very fact that some conscious creatures are treated as a resource for other conscious creatures, I contend, should raise in us moral indignation, indeed the very same moral indignation. This is the deep motivation for defending animal rights.

5. An Objection

I have presented a moral framework and have argued that it grants expansive rights to conscious animals, but also that probably only mammals are conscious. I have concluded that probably only mammals possess the relevant rights. There are many objections that could be lodged against this view, and I cannot discuss all of them, or even many of them here. There is one particularly important objection that I do want to discuss: the objection that even a low probability of consciousness in an animal should suffice to grant it the relevant rights, because the moral costs of error in this area are gigantic. Even conceding that the COI hypothesis is the most plausible hypothesis about the neural correlates of consciousness, surely other hypotheses are still importantly plausible, and may have the consequence that certain animals which the COI hypothesis considers unconscious are in fact

conscious. The mere plausibility of this alternative hypothesis suffices, according to the objector, for granting the relevant rights to the animals considered conscious by it (see Singer 1975 for a similar consideration).

The objection essentially denies the inference from the claims that only conscious animals should be granted the relevant rights and that non-mammals are probably unconscious to the conclusion that non-mammals should probably not be granted these rights. As long as there is a non-negligible probability that an animal is conscious, we should grant it the rights of conscious animals, because doing so guarantees that we would not be unwittingly committing moral horrors. In reaction, I want to embrace the spirit of the objection but reject its letter (at least on one reading of what the letter is). My discussion will make use of certain toy examples that greatly oversimplifies the matter, but this will help to see what I think is right about the spirit of the objection and where its letter may go wrong.

First, however, we should keep in mind a distinction between two ways in which considerations of probability enter the picture. We may think of the picture developed in the previous section as a picture tailored to an idealized scenario, in which our credence in the COI hypothesis is 100%. As a matter of fact, it is not. My own credence in it is probably around 55%. So my credence in the conclusions based on it cannot exceed 55%. Conversely, since my credence in the very liberal hypotheses about the neural correlates of consciousness, such as Zeki's, hovers around 5%, I can be upward 5% credent that animals as simple as oysters ought to be treated as ends and not mere means. So this is one way in which a more attenuated picture is forced on us by consideration of probability. However, the objection before us pushes beyond this. The objector claims rather that if I have 5% credence in oysters being conscious, I ought to have 100% credence that oysters ought to be treated as ends and not mere means.

Put this way, I think the objector's claim is too strong to be plausible. According to panexperientialists, everything in the universe has a degree of phenomenal consciousness, if only a very small one. Since their position is coherent,

and some of the arguments for it are not altogether ludicrous (Chalmers 1996), I may have 1% credence that vegetables are phenomenally conscious as well. Surely, however, I ought not to treat vegetables as ends in themselves on the grounds that panexperientialism is coherent and non-ludicrous. I take this to be a *reductio* of the strong claim that any non-negligible probability of consciousness entails treating as an end: the result of adopting that claim and recognizing that panexperientialism generates a non-negligible probability that everything is conscious is to treat everything as an end, which *ceteris paribus* entails that we ought to avoid eating anything.

It would, of course, be entirely arbitrary to declare that although a non-negligible probability of consciousness does not secure rights, a probability of 8%, or 12%, or some such figure, does. More generally, there is something wrong about the idea that there is a cut-off point beyond which credence in an animal's consciousness should entail 100% credence in the animal having the rights of conscious animals. A more sensible model would embrace a non-trivial function from credence in consciousness to credence in rights. The dogmatic model we have worked with thus far matched 12% credence in an animal's being conscious to 12% in it deserving the rights of conscious animals. The objector's model matches 12% credence in the animal's being conscious to 100% credence in it deserving the rights. Neither is plausible. It is much more reasonable, it seems to me, to adopt a function that, say, matches 12% credence in consciousness to 40% credence in rights. Obviously, we cannot expect that there be objective facts about the 'right' function. But I hope that this illustration makes clear what I think is right about the spirit of the objection before us. What is right about it is that the dogmatic function from credence in consciousness to credence in rights is wrongheaded: the moral cost of making a moral error is such that a more 'proactive' function should be embraced.

For example, suppose my credence in chickens having consciousness is 34%. The 'right' function may match a credence of 86% in chickens deserving to be treated as ends and not mere means. This means, roughly, that I should treat

chickens as ends in 86% of my dealings with them, or roughly in six out of seven interactions.²⁶

One way to think of the 'right' function issue is in terms of the stability and dominance of one's disposition to treat animals as ends and not mere means. As noted, above, the more stable and dominant her disposition, the more morally commendable the agent is (modulo the mentioned worries about puritanism). We can incorporate the considerations we have been discussing in the present section into this framework: *mutatis mutandis*, an agent who treats an animal in whose consciousness she has 12% credence as an end and not mere means 40% of time is more commendable than an agent who treats such an animal as an end 30% of the time and less commendable than one who treats it as an end 50% of the time. Likewise, an agent who treats an animal in whose consciousness she has 12% credence as an end and not mere means 40% of time is *mutatis mutandis* more commendable than an agent who treats an animal in whose consciousness she has 17% credence as an end and not mere means 40% of the time and less commendable than one who treats in such a way an animal in whose consciousness she has 7% credence. This is because of variation in the stability and dominance of the key disposition in these various agents.

As warned above, this discussion makes free use of oversimplifying toy examples. The reality of moral conduct is not as clean as the assignment of precise figures in the above discussion suggests. I have adopted this way of discussing the objection that is the subject matter of this section merely as an expository device, by way showing what I think is right about the objection and what I think is wrong. It is right that the mere probability that an animal is conscious should importantly boost its claim to having moral rights. It is wrong that any non-negligible (or arbitrary) probability secures those rights absolutely.

6. Conclusion

The history of moral progress is the history of widening the circle of moral worth, whereby each group grants more 'others' the same rights they enjoy themselves. Certainly in the West we have witnessed this process in the form of dehumanizing less and less (first with respect to other ethnicities, then other races, and so on and so forth). Currently on the agenda in the West is the widening of the circle to people of different sexual orientation, and since history has nowhere to march but forward, this issue will most certainly be eventually settled in favor of equal rights across orientations. But where does the process end? Where is the point beyond which there is no need to progress? In short, what is the end of moral history? It seems to me that the circle of the conscious is the widest circle of moral worth, and moral history will come to an end when all conscious creatures will enjoy equal rights. My own more specific suggestion, for which I have argued here, is that this point will arrive when all conscious animals will be treated as ends in themselves and not merely as means.²⁷

References

- Amsterdam, B. 1972. "Mirror self-image reactions before age two." *Developmental Psychobiology* 5: 297-305.
- Arico, A.J., B.Fiala, R.F. Goldberg, and S.B Nichols. **Forthcoming**. "The Folk Psychology of Consciousness."
- Armstrong, D. M. 1968. *A Materialist Theory of the Mind*. New York: Humanities Press.
- Baars, B.J. 1988. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge UP.
- Block, N.J. 1995. "On a Confusion About the Function of Consciousness." *Behavioral and Brain Sciences* 18: 227-247.
- Block, N.J. 2007. "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience." *Behavioural and Brain Sciences* 30: 481-499.

- Brentano, F. 1874. *Psychology from Empirical Standpoint*. Edited by O. Kraus, editor of English edition L. L. McAlister, translation A. C. Rancurello, D. B. Terrell, and L. L. McAlister. London: Routledge and Kegan Paul, 1973.
- Carruthers, P. 1999. "Sympathy and Subjectivity." *Australasian Journal of Philosophy* 77: 465-482.
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge UP.
- Carruthers, P. 2004.
- Caston, V. 2002. "Aristotle on Consciousness." *Mind* 111: 751-815.
- Chalmers, D. J. 1995. "Facing Up to the Problem of Consciousness." *Journal of Consciousness Studies* 2: 200-219.
- Chalmers, D. J. 1996. *The Conscious Mind*. Oxford and New York: Oxford UP.
- Choi, S. 2008. "Dispositional Properties and Counterfactual Conditionals." *Mind* 117: 795-841.
- Churchland, P.S. 1984. *Matter and Consciousness*. Cambridge MA: MIT Press.
- Dehaene, S., C. Sergent, and J.-P. Changeux 2003. "A Neuronal Network Model Linking Subjective Reports and Objective Physiological Data During Conscious Perception." *Proceedings of the National Academy of Science USA* 100: 8520-8525.
- Dennett, D.C. 1991. *Consciousness Explained*. Boston: Little, Brown, & Co.
- Dretske, F. I. 1995. *Naturalizing the Mind*. Cambridge MA: MIT Press.
- Fara, M. 2005. "Dispositions and Habituals." *Noûs* 39: 43-82.
- Flavell, J. H. 1979. "Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry." *American Psychologist* 34: 906-911.
- Fodor, J.A. 1992. "A Theory of the Child's Theory of Mind." *Cognition* 44: 283-296.
- Gallup, G.G. 1970. "Chimpanzees: Self-recognition." *Science* 167: 86-87.
- Gallup, G.G. 1975. "Towards an operational definition of self-awareness." In R. H. Tuttle (ed.), *Socioecology and the Psychology of Primates*. Berlin: Mouton de Gruyter.
- Gehring, W.J., R.T. Knight 2000. "Prefrontal-Cingulate Interactions in Action Monitoring." *Nature Neuroscience* 3: 516-520.
- Gopnik, A. and J. W. Astington 1988. "Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction." *Child Development* 59: 26-37.
- Gray, J.A. 2004. *Consciousness: Creeping up on the Hard Problem*. Oxford and New York: Oxford UP.
- Griffin, J. 1986. *Well-Being*. Oxford: Clarendon Press.
- Kriegel, U. 2003. "Consciousness as Intransitive Self-Consciousness: Two Views and an Argument." *Canadian Journal of Philosophy* 33: 103-132.
- Kriegel, U. 2007. "A Cross-Order Integration Hypothesis for the Neural Correlate of Consciousness." *Consciousness and Cognition* 16: 897-912.
- Kriegel, U. 2009. *Subjective Consciousness: A Self-Representational Theory*. Oxford: Oxford UP.

- Kriegel, U. and K.W. Williford 2006. *Self-Representational Approaches to Consciousness*. Cambridge MA: MIT Press.
- Lau, H.C. and R.E. Passingham 2006. "Relative Blindsight in Normal Observers and the Neural Correlate of Visual Consciousness." *Proceedings of the National Academy of Science USA* 103: 18763-18768.
- Leopold, D.A. and N.K. Logothetis 1996. "Activity Changes in Early Visual Cortex Reflects Monkeys' Percepts During Binocular Rivalry." *Nature* 379: 549-553.
- Lewis, D.K. 1997. "Finkish Dispositions." *Philosophical Quarterly* 47: 143–158.
- Lurz, R. 1999. "Animal Consciousness." *Journal of Philosophical Research* 24: 149-168.
- Lycan, W. 1996. *Consciousness and Experience*. Cambridge MA: MIT Press.
- Molnar, G. 1999. "Are Dispositions Reducible?" *Philosophical Quarterly* 49: 1–17.
- Nordgreen, J., J. Garner, A. Janczak, B. Ranheim, W. Muir, T. Horsberg. **Forthcoming**. "Thermonociception in fish: Effects of two different doses of morphine on thermal threshold and post-test behaviour in goldfish (*Carassius auratus*)." *Applied Animal Behaviour Science*.
- Onishi, K.H. and R. Baillargeon 2005. "Do 15-Month-Old Infants Understand False Beliefs?" *Science* 308: 255-258.
- Perner, J., S.R. Leekam, and H. Wimmer 1987. "Three-year-olds' difficulty with false belief: The case for a conceptual deficit." *British Journal of Developmental Psychology* 5: 125- 137.
- Plotnik, J.M., F.B.M. de Waal, and D. Reiss 2006. "Self-Recognition in an Asian Elephant." *Proceedings of the National Academy of Sciences* 103: 17053-17057.
- Povinelli D.J., A.B. Rulf, and D.T. Bierschwale. 1994. "Absence of knowledge attribution and self-recognition in young chimpanzees (*Pan troglodytes*)." *Journal of Comparative Psychology* 108: 74–80.
- Quine, W.V.O. 1960. *Word and Object*. Cambridge MA: MIT Press.
- Reiss D. and L. Marino 2001. "Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence." *Proceedings of the National Academy of Sciences USA* 98: 5937-5942.
- Rosenthal, D. M. 1986. "Two Concepts of Consciousness." *Philosophical Studies* 94: 329-359.
- Rosenthal, D. M. 1990. "A Theory of Consciousness." ZiF Technical Report 40, Bielfield, Germany. Reprinted in N.J. Block, O. Flanagan, and G. Guzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press, 1997.
- Rosenthal, D. M. 1993. "Thinking that One Thinks." In M. Davies and G. W. Humphreys (eds.), *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Rosenthal, D. M. 2000. "Consciousness and Metacognition." In D. Sperber (ed.), *Metarepresentation*. Oxford: Oxford UP.
- Rosenthal, D. M. 2002. "Explaining Consciousness." In D. J. Chalmers (ed.), *Philosophy of Mind*. Oxford and New York: Oxford UP.
- Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson & Co.

- Sahraie, A., L. Weiskrantz, J. L. Barbur, A. Simmons, S. C. R. Williams, and M. J. Brammer 1997. "Pattern of neuronal activity associated with conscious and unconscious processing of a visual signal." *Proceedings of the National Academy of Sciences USA* 94: 9406-9411.
- Singer, P. 1975. *Animal Liberation*. New York: Random House.
- Song, H. and R. Baillargeon 2008. "Infants' Reasoning about Others' False Perceptions." *Developmental Psychology* 44: 1789-1795.
- Suarez S. D. and G. G. Gallup 1981. "Self-recognition in chimpanzees and orangutans, but not gorillas." *Journal of Human Evolution* 10: 175-188.
- Tomasello, M. and J. Call. 2005. "Do chimpanzees know what others see – or only what they are looking at?" In M. Nudds and S. Hurley (eds.), *Rational Animals*. Oxford: Oxford UP.
- Tye, M. 2000. *Consciousness, Color, and Content*. Cambridge MA: MIT Press.
- von der Malsburg, C. 1981. "The Correlation Theory of Brain Function." Technical Report 81-2, Max-Planck-Institute for Biophysical Chemistry, Gottingen.
- Weiskrantz, L. 1986. *Blindsight: A Case Study and Implication*. Oxford: Oxford UP.
- Williford, K.W. 2006. "The Self-Representational Structure of Consciousness." In Kriegel and Co-author 2006.
- Zeki, S. and A. Bartels 1999. "Toward a Theory of Visual Consciousness." *Consciousness and Cognition* 8: 225-259.

¹ Note that although Chalmers draws the distinction in the distinction in the context of defending a dualist theory of consciousness, the mere drawing of the distinction does not commit us to dualism. First of all, it may well be that the phenomenal property picked out by the phenomenal conception may be identical to some physical/neural property. Secondly, the psychological/phenomenological distinction concerns characterizations and conceptions, not properties. It may well be that the phenomenological characterization/conception picks out what is in fact a psychological property, it just does not characterize it *as* a psychological property.

² At least we can *stipulate* that this is so. It may be that the case in the text is underdescribed to guarantee that this is the case, but the story could be filled out so that it does.

³ To repeat, this does not commit us to dualism, for the reasons brought up in [Note 1](#).

⁴ Different versions of consequentialism play out in different ways, of course, and there are some versions that may focus on pleasure or on pain exclusively (e.g., Bentham's original version, which was formulated in terms of pleasure/happiness only).

⁵ Sometimes the term 'person' is used to refer to any subject of conscious experience, or any subject of moral worth, whereas on other occasions it is used to refer only to human subjects, or subjects with relatively sophisticated inner life. I use it here in the former sense, and add 'or at least of subjects' for the benefit of those who insist on using it the latter way. This is just a terminological matter.

⁶ This is not to say that Singer is committed to the entire picture of the nature of pain and pleasure that we have developed thus far.

⁷ More often Rosenthal actually formulates the principle somewhat differently: as the claim that ‘conscious states are states we are conscious of.’ But this unnecessarily generates an air of circularity. The circularity may be merely apparent, if one accepts Rosenthal’s (1986) distinction between *transitive* and *intransitive consciousness*, as well as his claim that the former can occur in the absence of the latter (a claim I argue against in Kriegel 2004 and 2009). However, to avoid even the appearance of circularity, it is wiser to frame the principle in terms of awareness.

⁸ In Kriegel 2004 and Kriegel 2009 Ch.7 I propose an account of a non-reflective, relatively unimposing kind of awareness which is typically directed at our conscious states.

⁹ Note that on such functionalist views, there is no genuine distinction between the phenomenal and the psychological. This is, of course, a reason for objecting to such views (Block 1995, Chalmers 1995). It should be stressed, though, that the ‘if’ claims in the text are meant to hold metaphysically, not merely nomologically. Chalmers (1996), for example, holds that the right functional role is nomologically sufficient, but metaphysically insufficient, for phenomenal consciousness.

¹⁰ It is nevertheless possible to take psychological pain to constitute *evidence* for the presence of phenomenological pain. This more indirect use of psychological pain will be discussed later.

¹¹ We will see later why the beliefs must be false and the representations non-veridical (i.e., be misrepresentations).

¹² Going into the experimental details will distract us here, but let me just report the sociological observation that Povinelli’s is distinctly a minority view within the scientific community, and that the consensus still favors Tomasello, and quite heavily.

¹³ For the findings on mirror self-recognition in human infants and toddlers, see Amsterdam 1972. For its lack in gorillas, see Suarez and Gallup 1981. For the surprising successes of dolphins and Asian elephants, see (respectively) Reiss and Marino 2001 and Plotnik et al. 2006. Gallup was the first to conduct mirror self-recognition experiments (Gallup 1970), and later offered them as capturing an operational definition of self-awareness (Gallup 1975).

¹⁴ A reductive account might *identify* phenomenological pain with inner awareness of a psychological pain. But even if we do not take a reductive view of the matter, phenomenological pain would *correlate* with inner awareness of psychological pain rather than with psychological pain on its own.

¹⁵ The term ‘seriously’ is problematic, inasmuch being taken seriously comes in degrees. Some philosophers – eliminativists – have argued that *nobody* is conscious, not even humans (Churchland 1984). This has always been a marginal view, however, in a way Povinelli’s view has not. I recognize, however, that there may be an element of artificiality in deciding what we should take as unquestionable cases of consciousness. Perhaps all apes could be taken as unquestionably conscious.

¹⁶ One occasionally finds scientists declaring that an invertebrate or fish can feel pain, but this is invariably based on inferential leaps that assume a tighter connection between psychological and phenomenological pain than is consistent with the transitivity principle (see, e.g., Nordgreen et al. forthcoming).

¹⁷ Note that in claiming that only mammals are conscious, we are not committing to it being the case that *all* mammals are.

¹⁸ Most or all of the advantages of the non-consequentialist framework I am adopting over the consequentialist framework may apply to the straight Kantian framework and be unaffected by the virtue-ethical twist. I introduce the virtue-ethical twist simply because the resulting precept (i) is what I actually believe and (ii) has never been formulated (to my knowledge).

¹⁹ This is mainly because of a number of counter-examples involving so-called finks, masks, and antidotes. The primitivist account claims that dispositions are primitive and cannot be analyzed in terms of more basic notions.

²⁰ It is an intriguing question what this would imply in practice. One possibility is to force ourselves to treat conscious creatures as mere means on occasion. Just how often 'on occasion' should be ought to be determined by psychological research on the forming of the dour and intolerant attitude of which I spoke. It may well be that such research will discover that my observation is simply wrong (perhaps it is based on the wrong anecdotes), or that it can be avoided by simpler means. Whatever such research turns up would have to be integrated into the Kantian virtue-ethical stance I describe here.

²¹ These two views are, moreover, clearly compatible. Unfortunately, and somewhat surprisingly, there is relatively little work on the analysis of treating someone as an end, even among Kant scholars. Moreover, it would seem that Kant himself had little to offer on this matter, so that many interpretations are in principle open (see Griffin 1986). But the Kantian proponent of animal rights could just suggest here that whenever some consensus emerges regarding the correct analysis of this matter, it could be plugged into (CF+).

²² At the same time, note that (CF+), like (CT+), does not prohibit eating, say, cows altogether: a cow dead of natural causes, since it is no longer conscious, has no rights, and can be eaten. (This is, by the way, the current practice in Bhutan.)

²³ Again, it is important to keep in mind that this typically conscious animals brought up today for consumption of their products live such a horrendous life of suffering that (CT+) does in fact prohibit consumption of their products.

²⁴ A qualification is necessary here. My claim is that using the products of these unconscious animals is not intrinsically wrong. But there may be ways in which it is instrumentally wrong. For example, gratuitous torturing of an unconscious animal that elicits in us the intuitive ascription of consciousness is sure to do harm to our character, and therefore such torturing is morally wrong, though only instrumentally. Perhaps the treatment of geese in the context of *foie gras* production is a good example: it is not intrinsically wrong, on my view, since geese are probably unconscious, but it is instrumentally wrong, in that it desensitizes us to what would be cruel treatment if geese were conscious.

²⁵ At the same time, if a person grows a cow as a pet and a companion, and milks her only when she is naturally in need of milking, it is morally permissible for him to consume her milk; likewise if he uses her skin for clothing when she dies a natural death – or her flesh for food.

²⁶ One practical interpretation of this may be that one ought to eat chicken at most once a week if one has a 34% credence in their being conscious.

²⁷ For very helpful comments on a previous draft, I would like to thank Sebastian Leugger, Krysta Vollbrecht, and especially Stephen Biggs. For important conversations, I would like to thank Benji Kozuch, David Rosenthal, Mark Timmons, and especially Krysta Vollbrecht.