Moral status, luck, and modal capacities: debating Shelly Kagan

Harry R. Lloyd

Department of Philosophy,
Yale University,
P.O. Box 208306,
New Haven, CT 06520-8306,
USA.

harry.lloyd@yale.edu

ABSTRACT: Shelly Kagan has recently defended the view that it is morally worse for a human being to suffer some harm than it is for a lower animal (such as a dog or a cow) to suffer a harm that is equally severe (ceteris paribus). In this paper, I argue that this view receives rather less support from our intuitions than one might at first suppose. According to Kagan, moreover, an individual's moral status depends partly upon her 'modal capacities.' In this paper, I argue that the most natural strategy for justifying Kagan's theory faces some important challenges. More generally, I argue that philosophers who wish to defend the view that human beings have a higher moral status than that of the lower animals face a dilemma. Either their theory of moral status will imply (unacceptably) that some severely cognitively impaired human beings have a significantly lower moral status than that of typical human beings, or these philosophers will be forced to ground moral status in a set of properties so far removed from a subject's actual capacities that it will become difficult to see why these kinds of properties should have such moral importance.

I. Introduction

In *How to Count Animals, more or less*,[1] and in an earlier lecture printed in this journal,[2] Shelly Kagan argues that the lower animals (such as cows and dogs) have a lesser 'moral status' than most human beings – call this view *hierarchy*. In particular, one of the views that Kagan defends is the view that, *ceteris paribus*, and even putting aside concerns about distributive justice, it is morally worse for a typical human being to suffer some harm than it is for a lower animal to suffer a harm that is equally severe – and mutatis mutandis for benefits (HCA, §4.3) – call this view *simple axiological hierarchy*.[3] Kagan thinks that our intuitions concerning a certain type of thought experiment support this view; I challenge that claim in §II of this paper.

Axiological hierarchy concerns the *goodness of outcomes*. One way in which one could incorporate hierarchy into one's theory of *right action* would be to combine axiological hierarchy with consequentialism. Alternatively, a deontologist could incorporate hierarchy into her theory of right action (regardless of whether or not she endorses axiological hierarchy) by maintaining that whilst (at least most) human beings have certain strong *natural rights* against being harmed or killed, the natural rights of lower animals are, at the very least, more easily permissibly infringed – call this view *rights-focused hierarchy*.[4] A rights-focused hierarchy theorist might, for example, hold the view that for any given positive value h, it is morally impermissible to inflict h utils worth of pain upon a victim, unless doing so will generate at least mh utils worth of pleasure, where the value of the multiplier m is proportional to the moral status of the victim of the harm (HCA, §9.2).

One possible alternative to (both types of) hierarchy is the view that lower animals lack any kind of moral status whatsoever – call this view *chauvinism*. Immanuel Kant is a famous example of a chauvinist, since he argues that one ought to avoid cruelty towards lower animals only because this has a beneficial effect upon one's own character, making it less likely that one will be cruel to human beings in the future.[5] Kantian chauvinism is highly implausible: even if one could find a way to gratuitously torture lower animals without this having any deleterious effect upon one's moral character, it would nonetheless be egregiously morally wrong.

Another possible alternative to hierarchy is the view that all or almost all animals (perhaps with some exceptions, like sponges) have the same moral status – call this view *unitarianism*.[6] Chauvinists and hierarchy theorists are both *anti-unitarians*. I believe that unitarianism is correct, and one of my key aims in this paper is to draw attention to a dilemma that confronts anti-unitarian theorists like Kagan.

Here's the dilemma faced by anti-unitarian theorists: either (*first horn*) their theory will imply (unacceptably) that severely impaired human beings with similar cognitive capacities to the lower animals have a significantly lower moral status than that of typical human beings, or (*second horn*) these theorists will be forced to ground moral status in a set of properties so far removed from a subject's actual capacities that it will become difficult to see why these kinds of properties should have such moral importance. I illustrate the anti-unitarian's dilemma in §III of this paper: theories of moral status that appeal exclusively to features like cognitive and/or emotional capacities are impaled upon the first horn, whereas theories of moral status that appeal to species membership are impaled upon the second.

In §§IV-VI, I discuss Kagan's theory of moral status. I begin, in §IV, by summarising some of the theory's attractive features. I argue in §§V-VI, however, that the most natural strategy for extricating this theory from the second horn of the anti-unitarian's dilemma faces some important challenges. In §VII, I argue that unitarianism avoids both horns of the anti-unitarian's dilemma. In §VIII, I conclude.

II. Our intuitions about hierarchy

Before I discuss the anti-unitarian's dilemma, I want to briefly discuss one of the thought experiments that (partially) motivates Kagan's hierarchical theory. In attempting to establish that simple axiological hierarchy is intuitively plausible, Kagan – echoing an earlier argument made by Jean Kazez[7] – suggests that we should consider a case in which a human being and a

mouse are suffering from equally painful toothaches (HCA, p. 99). Kagan reports the intuition that it would be more important, morally speaking, to avert the human toothache than it would be to avert the mouse toothache.[8]

Speaking here as a card-carrying consequentialist, I agree with Kagan that averting the human toothache is almost certainly more important than averting the mouse toothache. But I disagree that this stands as evidence in favour of axiological hierarchy. In virtue of our advanced cognitive capacities, we human beings can, and often do, engage in certain highly valuable projects that aren't open to mice. We build friendships, acquire systematic knowledge, and express ourselves, all in ways that go far beyond what any mouse could ever imagine.[9] But: we cannot do any of these things particularly well when we have a toothache![10] Thus, I claim that we form our intuitive reaction to the Kagan-Kazez thought experiment based on the justified assumption that there is more welfare at stake for the human being than there is for the mouse.

To try to get some proper intuitions about axiological hierarchy, we might instead consider a case in which a human being and a dog suffer equally painful cases of stubbing their toe. Crucially, we should suppose that the human being was walking along doing nothing in particular, that the pains quickly subside, that they were not anticipated, and that they are immediately forgotten.[11] In evaluating this case, I for one am inclined to intuit that the human pain is no more important, morally speaking, than the dog's pain.[12] If the dog's pain were to be slightly higher, then I would prioritise averting it over averting the human being's pain.

My aim in presenting this toe-stubbing case is to suggest that thought experiments which involve comparing a harm to a human against a harm to an animal provide considerably less support to axiological hierarchy than one might at first suppose. It would be wrong to claim, however, that my intuitive reaction to the toe-stubbing case is inconsistent with every potential form of simple axiological hierarchy.[13] According to one version of simple axiological hierarchy, for example, an individual's moral status affects the moral disvalue of harming her only for harms that exceed a certain level of severity, which cannot be met by short-lived harms – like toe-stubbings – that are immediately forgotten. Likewise, consider the (rather less *ad hoc*) hierarchical view that the moral disvalue D of a harm that inflicts h utils worth of pain is given by $D = (1+\varepsilon)^{mh}$, where ε is small and positive, and m is proportional to the moral status of the victim. D will be insensitive to m when h is small, but much more sensitive to m when h is large. Thus, the moral disvalues of the dog's toe stubbing and of the human's toe stubbing (in the context of my thought experiment) will, according to this theory, be close to identical even if the cross-species difference in m is large, since the common h is relatively small.

Some people might disagree with my intuitive response to the toe-stubbing thought experiment, instead intuiting that the dog's pain matters less than the human's pain. Perhaps, however, we can attempt a debunking response to that kind of intuition. Philosophers and cognitive scientists since Hume and Adam Smith have recognised the important role that *empathy* plays in the formation of our moral judgements.[14] Perhaps axiologically hierarchical intuitions are the product of deficiencies in our empathetic faculty. I, for one, find it harder to imagine what it is like to be a dog than to imagine what it is like to be another human being. Hence, I often find it easier to empathise with human beings than to empathise with dogs. This

looks more like evidence of a deficient empathetic faculty, prone to in-group and similarity biases, than it looks like evidence of dog pain mattering less than human pain.[15]

Our lack of empathy towards the lower animals might be partly caused by the fact that our society's pervasive use of animals for food, clothing, and medical testing encourages an *instrumental* attitude towards animals: a recent study of Italian veterinary students found evidence of a link between instrumental perceptions of the lower animals and reduced empathy for them (although the direction of causation is still somewhat unclear).[16] Our lack of empathy for the lower animals might also be partly caused by the fact that most of us tend to witness animal suffering much less frequently than we witness human suffering. As Mengzi remarks (*Mencius* 1A7), "as for the relation of gentlemen to birds and beasts … if they hear their cries, they cannot bear to eat their flesh. Hence, gentlemen keep their distance from the kitchen [where animals are slaughtered to be eaten]."[17]

Still, many people's intuitions probably support rights-focused hierarchy more robustly than they support simple axiological hierarchy. Many people probably intuit both (a) that it is morally impermissible to harvest one healthy human being's organs without their permission in order to save five others, and also (b) that it is at worst morally permissible, and perhaps even morally obligatory, to harvest one healthy rabbit's organs in order to save five other rabbits (HCA, p. 192).[18] Of course, one might attempt a debunking response to these kinds of intuitions too. However, I want to argue in the remainder of this paper that rights-focused hierarchy is difficult to sustain even if we concede that it is prima facie intuitively plausible.

III. The naïve actual-capacities theory, and speciesism

On what grounds might one try to sustain hierarchy about moral status? According to one initially attractive proposal, an individual's moral status depends upon her cognitive and/or emotional capacities, with lower animals having a moral status considerably lower than that of a typical human being – call this the *(naïve) actual-capacities theory*. It has struck several critics of the simple-axiological version of the naïve actual-capacities theory as implausible to suppose, however, that an individual's cognitive and/or emotional capacities are the sort of things that could bear upon the moral (dis)value of harming or benefitting her (HCA, §4.4). These critics ask: why should the fact that human beings typically possess advanced cognitive capacities constitute a reason for thinking that it is worse to harm a typical human being than to inflict an *equally severe* harm upon a lower animal? I think that many people, even some who are initially attracted to simple axiological hierarchy, will feel greater intuitive unease about assigning axiological importance to an individual's cognitive and/or emotional capacities than they would feel about assigning axiological importance to, say, the durations of harms, or to whether an individual is worse off than others through no fault of her own.

On this score the prospects for the rights-focused version of actual-capacities hierarchy appear to be somewhat brighter. Philosophers who believe in the existence of natural rights often ground these rights in the fact that typical human beings are 'autonomous persons', cognitively capable of practical reasoning about personalised life-plans and projects (HCA, §8.2, EK, pp. 252-8, 473-85). The rough idea here is that the purpose of natural rights is to, in some sense, 'protect' the autonomy of such persons. The rights-focused actual-capacities hierarchy theorist can now point out, however, that the lower animals are much less capable than typical human beings are at engaging in this kind of practical reasoning, precisely because they have less-sophisticated cognitive capacities. According to this view, the robustness of

one's natural rights ought to depend upon one's degree of autonomy,[19] and one's degree of autonomy in fact depends upon (a certain subset of) one's cognitive capacities.

Both versions of the naïve actual-capacities theory are, however, impaled upon the first horn of the anti-unitarian's dilemma. The naïve actual-capacities theory implies that some severely cognitively impaired human beings (whose cognitive and emotional capacities are similar to those of the lower animals) have a comparable moral status to that of the lower animals, rather than to that of typical human beings. I think that many people would intuit, to the contrary, that severely cognitively impaired human beings have the same rights as those held by typical human beings, at least insofar as it is no more permissible to harvest the organs of one severely cognitively impaired human in order to save *n* other severely impaired humans than it is permissible to kill one typical human in order to save *n* other typical humans. And even those people who do not share this intuition will probably wish to claim that severely impaired human beings have rights that are more closely comparable to the rights of typical human beings than they are to the rights of lower animals.[20] Any version of the naïve actual-capacities theory that puts a significant moral distance between typical human beings and the lower animals is impaled upon the first horn of the anti-unitarian's dilemma (see §I above).[21]

One way to avoid the first horn of the anti-unitarian's dilemma is to adopt the view that an individual's moral status depends upon the typical capacities of a member of her species – call this view *speciesism*. This approach too, however, faces some problems. First of all, it is difficult to see why the typical capacities of a member of one's species should have any significant moral importance. Why is this as important as – or more important than – an individual's own capacities? In any case, why is species (on any candidate definition, of which there are many[22]) the right taxon for us to refer to here?[23] Speciesism is impaled upon the second horn of the anti-unitarian's dilemma.

Speciesism is also vulnerable to counterexamples. For sake of argument, suppose that we adopt Ernst Mayr's well-known 'biological' conception of *a species* as: a group of reproductively isolated individuals, capable of interbreeding to produce fertile offspring.[24] Under that assumption, here's a counterexample to speciesism: suppose that I carry out a worldwide gene-editing campaign, splitting humanity into two species. A-humans can breed to produce fertile offspring with other A-humans; likewise for B-humans; and this is the full extent of my genetic meddling. Now, suppose that because of an astronomically unlikely case of coincident de novo genetic mutations, *all* of the next generation of A-humans are born with a severe and heritable cognitive deficiency. Suppose that a small proportion of the next generation of B-humans are also born with the deficiency. Speciesism implies that the impaired A-humans have a lower moral status than the impaired B-humans. This implication is implausible, and should be rejected.[25] (Note that this particular counterexample cannot be sidestepped by a theory in which (a) membership of a sophisticated species and (b) personal possession of sophisticated capacities each count as sufficient but unnecessary conditions for possessing full moral status.)

IV. Kagan's hierarchical theory

Kagan's hierarchical theory of moral status can be construed as an attempt to avoid the common counterexamples to both speciesism and the actual-capacities theory. According to Kagan's theory, an individual's moral status depends not only upon her actual capacities, but also upon (a) her *modal capacities* – the capacities that she *could have had* (HCA, §5.5) – and (b) her

potential capacities – the capacities that she stands to have in the future (HCA, §5.4). A *modal person* is someone who could have had the capacities necessary for things like rationality, self-consciousness, and practical reasoning – the capacities that are constitutive of *personhood*. To make Kagan's theory work, we will have to adopt a conception of modal capacities, modal personhood, and 'could have had' that comes in degrees (HCA, p. 140) – and likewise for potential capacities and 'stands to have' (HCA, pp. 136-7). The strength of an individual's modal personhood, for example, will depend upon the degree of similarity (for some suitable definition of 'similarity', which I'll leave implicit in this paper) between (a) the actual world and (b) the nearest possible world(s) in which that individual is a person.

Kagan argues that his theory of hierarchy can be used to defend the view that severely cognitively impaired human beings have a moral status that is closer to that of a typical human being than it is to that of a lower animal (HCA, §6.3). (*i*) Somebody who suffers brain damage in the womb or in a car accident is a strong modal person, because she could have (to this day) been a person, if only she had not suffered her accident. (*ii*) Likewise, someone who is cognitively impaired because of a genetic condition is a strong modal person, because she could have developed into a person, if only her genome had been *slightly* different. Kagan's theory thus appears to improve upon speciesism, because it does not imply that A-humans (see §III above) have the same moral status as the lower animals. Were it not for their 'unlikely' genetic mutations, the A-humans could have been persons. Hence, the A-humans have a moral status that is similar to that of typical human beings, and equal to that of any impaired B-humans.

Argument (*ii*) here is controversial, because it requires us to reject the view that no individual could have been born with a different genome to the one that she was actually born with. In other words, it requires us to reject the view that having been born with the genome that one was actually born with is an *essential property* of any individual – call this view *genetic essentialism*. Kagan tells us that he is minded to reject genetic essentialism, although he does not explain why (HCA, p. 161). As it happens, several moral philosophers have recently advanced arguments against genetic essentialism.[26] I obviously cannot hope to evaluate the merits of those arguments in this paper. Fortunately, neither do I need to: my intention in this paper is to show that *even if* genetic essentialism is false, and so Kagan's theory avoids the first horn of the anti-unitarian's dilemma, Kagan's theory is nonetheless impaled upon the second horn of the dilemma (see §§V-VI below).

In contrast to the two examples involving cognitively impaired human beings that we have just been discussing, consider the example of my dog, Fido. In certain possible worlds, a genetic therapy conferring advanced cognitive capacities upon dogs has already achieved widespread adoption. Many of us will be inclined to say that in these worlds, Fido would, in all likelihood, have become a person.[27] Nonetheless, because these worlds are very different from the actual world, Fido is only a very weak modal person (HCA, p. 140). Likewise, only very *extensive* changes to Fido's genome at conception could have produced an individual that would have gone on to develop the advanced cognitive capacities of a human person. Thus, it would be implausible to claim that Fido could have been born with the intrinsic potential to become a person, and so Kagan's theory of hierarchy implies that Fido has a lower moral status than that of human beings.

My discussion of Kagan's theory has thus-far focused upon his appeal to modal capacities. Kagan's appeal to potential capacities, however, is also important. Without this

appeal, Kagan's theory would implausibly imply that new-born human infants have a moral status comparable to that of the lower animals, rather than to that of typical adult human beings. I think that many people would intuit, to contrary, that new-born human infants have rights identical to those held by typical human beings, at least insofar as it is no more permissible to harvest the organs of one new-born human infant in order to save n other new-born human infants than it is to kill one typical human adult in order to save n other typical human adults. And even those people who do not share this intuition will probably wish to claim that new-born human infants have rights that are more closely comparable to the rights of typical human adults than they are to the rights of lower animals.

The plausibility of Kagan's theory of hierarchy depends, then, upon all three of his appeals to (a) actual capacities, (b) potential capacities, and (c) modal capacities. If Kagan cannot justify (b), then he cannot justify why new-born infants have a moral status similar to that of typical human adults. If he cannot justify (c), then he cannot justify why severely cognitively impaired human beings have (regardless of whether their impairment is congenital, or caused by an accident) a moral status similar to that of typical human beings. Compared to theories of hierarchy that have previously been proposed by other philosophers, however, Kagan's appeal to modal capacities is much more novel than his appeal to potential capacities. For this reason, my critique of Kagan's theory in §§V-VI of this paper will focus on the former, rather than the later.[28]

V. Justifying the appeal to modal capacities?

Suppose that one accepts the view that an individual's cognitive and/or emotional capacities are the sorts of things that might bear upon her moral status.[29] Why, we might then ask, should an individual's modal capacities, as well as her actual capacities, also affect her moral status? Without some kind of answer to this question, I think that many people, even amongst those who are initially attracted to hierarchy, will feel greater unease about assigning moral importance to an individual's modal capacities than they would feel about assigning moral importance to, say, the durations of harms, or to whether an individual is worse off than others through no fault of her own.[30]

The most promising strategy that I am aware of to respond to this question on Kagan's behalf draws upon considerations concerning *luck*. If moral status *were to not* in part depend upon modal capacities, then *bad luck* – suffering cognitive impairment as a result of an accident or a genetic defect, for example – would be capable of robbing a person of anything close to the moral status that she would otherwise have enjoyed, if only she had not been so unlucky. It seems plausible to suppose, moreover, that bad luck should not be capable of reducing a person's moral status this drastically.[31] This defence of Kagan's appeal to modal capacities relies on us accepting a certain *modal* analysis of 'bad luck'. Losing out on something, through no fault of one's own, counts as bad luck to the extent that one could have had it, in the degree-theoretic sense of 'could have had' that I introduced in §IV.[32]

Unfortunately (for Kagan), although this modal reading of 'bad luck' might well be faithful to a certain construal of that concept, I do not believe that it can do philosophical work in justifying Kagan's appeal to modal capacities. To see why this is so, consider the case of a child, whose parents lost their family fortune by making some bad investments on the stock market just before their child was born.[33] According to the modal conception of bad luck, this child is unluckier than a child whose family has been poor for generations. The possible worlds

in which the first child's parents chose to invest in more profitable companies are much more similar to the actual world than are any possible worlds in which the second child's parents are rich (on any tenable conception of the similarity relation between possible worlds). Even if in some sense the first child is 'unluckier' than the second, this is not a sense of the word 'luck' that should be invested with any moral significance. A *luck egalitarian* would want to treat both children equally, for example.

What makes both of these children unlucky in the morally relevant sense is that they are both unfortunate in their formative circumstances, and are both (obviously) not themselves responsible for this disadvantage.[34] Now, someone who wants to accept this claim and yet to resist the apparently absurd view that individuals who are born as cows are just as unlucky, in the morally relevant sense, as individuals who are born as severely impaired human beings[35] might wish to argue that what counts as a 'misfortune' for an impaired human being doesn't count as a misfortune for a cow, where 'fortune' is to be understood as a technical term that refers to how well an individual's life goes, or how she flourishes (EK, p. 145).[36] Nonetheless, anyone who tries to use this manoeuvre to preserve Kagan's appeal to modal capacities will merely be kicking the can down the road. If the strength of an individual's misfortune in failing to have certain capacities is to be equal to the modal strength of her possession of those capacities, then misfortune will have to be defined modally, in terms of the capacities that the individual could have had. Yet this will leave our account of misfortune vulnerable to counterexamples along the same lines as 'two children' counterexample (involving the recently impoverished parents) that I raised as an objection to the modal conception of bad luck.

My critique in this section has been strongly coloured by the assumption that the modal capacities hierarchy theorist must adopt a conception of 'could have been' that comes in degrees. This is Kagan's position. But could the modal capacities theorist instead adopt a *binary* conception of 'could have been', with modal personhood being something that an individual either possesses or does not? A thought experiment will help to express my reservations.

Consider a world in which an identity-preserving genetic therapy that somewhat increases the intelligence of dogs has achieved widespread adoption. This therapy is called *superdog therapy*, and a dog that has received it is called a *superdog*. Superdog therapy can only be delivered to infant dogs, and due to a mix-up at the vets, a certain dog Rover misses out on receiving it. The modal capacities theory implies that Rover is, in our imaginary world, a strong modal superdog (HCA, p. 143).

The point that I want to make here is that (a) the world as described in this thought experiment and (b) the world as it actually is sit at opposite ends of a continuum, along which possible worlds vary in terms of the degree of adoption of the superdog therapy. The modal capacities theorist who adopts a binary conception of modal superdoghood will have to say that at some watershed level of adoption, a marginal increase in uptake of the therapy will transform all those dogs who have not received the therapy from having low moral status to having near-superdog moral status. This view strikes me as implausible.

Before I move on, I want to draw attention to a further lacuna in the luck-based defence of Kagan's appeal to modal capacities. In his discussion of modal capacities, Kagan suggests that "modal status … appears to affect one's [moral] status in an *asymmetrical* way … what matters are the (relevant) possibilities that concern *higher* or more advanced sets of capacities,

not the various possibilities that concern *lower* or less developed ones" (HCA, p. 140). In other words, if a genetic fluke causes some dog Romulus to be born with the cognitive capacities of a typical human being, then the fact that Romulus simply 'got lucky' ought *not* to cause any substantial reduction in his moral status. That is, I agree, the intuitively plausible view: it would be a great injustice to afford Romulus natural rights any weaker than those we afford to human beings with the same cognitive capacities, and/or to regard Romulus' welfare as less important, morally speaking, than that of a human being with the same cognitive capacities.

But if we take the view that bad luck should not be capable of drastically reducing an individual's moral status, then shouldn't we also take the view that *good* luck should not be capable of drastically *increasing* an individual's moral status, as Kagan says happens in the Romulus case? As it stands, the luck-based defence of Kagan's appeal to modal capacities is thus at best incomplete. The luck-based defence does not yet tell us why we should accept the view that good luck can drastically change a person's moral status, whereas bad luck cannot.[37]

VI. Debunking the modal account of misfortune

In the last section, I argued that a modal conception of misfortune has counterintuitive implications in cases like that of the two poor children. Still, the motivating idea that it is worse for a person to narrowly miss out on something than never to have stood a chance of achieving it definitely has some initial intuitive appeal. I want to argue here, however, that our intuitions do not in fact support the view that narrowly missing out on something is *intrinsically* worse for a person than never standing a chance of achieving it. Rather, we intuit that narrowly missing out on something is often a dire state of affairs for a person in virtue of how it typically makes people *feel*.

When we narrowly miss out on something that we easily could have had – if, for example, we happen to be out buying coffee at the moment when Bill Gates pays a surprise visit to our company's office and starts handing out $100 bills[38] – we are likely to react with feelings of regret and annoyance. These sorts of feelings are unpleasant to experience, and they are also instrumentally disvaluable, insofar as they distract or demoralise us from pursuing our life plans. We would not, however, have experienced these feelings had Bill Gates simply never paid or intended to pay a surprise visit to the company office. Thus, narrowly missing out on something often makes us feel worse than never standing a chance of receiving it.[39]

But now suppose that a person narrowly misses out on Bill Gates' $100 bills without experiencing any of these negative emotions – suppose that she has an extremely stoical disposition, for example. Is narrowly missing out on Bill Gates' $100 bills any worse for this person than a state of affairs in which Bill Gates never visits her office, and thus in which she never stands a chance of receiving the $100 bills? I am inclined to suppose that it is not.

Hence, to the extent that we do sometimes intuit that narrowly missing out on something is intrinsically worse for a person than never standing a chance of achieving it, I am inclined to believe that our intuitive faculty is being corrupted by our intuitions about how narrowly missing out on something typically makes us feel. If only we were to think about the case more carefully, we would come to the conclusion that narrowly missing out on something is no more intrinsically unfortunate than never standing a chance of achieving it.[40]

VII. The advantages of unitarianism

According to the anti-unitarian's dilemma, hierarchical theories of moral status are stuck between a rock and a hard place. Complicated theories like Kagan's modal capacities theory manage to avoid counterexamples involving severely impaired human beings. But their complexity comes at a cost: it makes it difficult to see why the theorist's proposed determinants of moral status should have the degree of moral importance that is being ascribed to them. By contrast, it is easier for a rights-focused, actual-capacities theorist to explain why an individual's actual capacities should have an effect upon her moral status (see §III of this paper). But: the actual capacities theory is highly vulnerable to counterexamples involving severely impaired human beings. It remains to be seen whether any anti-unitarian theory of moral status can escape both horns of the anti-unitarian's dilemma.[41]

One reliable way to avoid both horns of the anti-unitarian's dilemma is to embrace unitarianism. Unitarian moral standing can be conditioned upon a binary property like sentience (the ability to experience pleasure and/or pain), and/or (bare) agency (the ability to have, and to wish to act upon, things like desires, wishes, and preferences) (see HCA, chapter 1). Prima facie, it is much more intuitively plausible to claim that properties like sentience and/or agency are the kinds of things that are intrinsically morally important and germane to an individual's moral status than it is to claim that modal capacities, say, are the kinds of things that are intrinsically morally important and germane to an individual's moral status.

Although sentience/agency-unitarianism avoids both horns of the anti-unitarian's dilemma, it does not imply that *all* human beings have moral standing. Since anencephalic infants (infants who are born without a cerebrum) lack the capacity for consciousness, sentience/agency-unitarianism implies that they lack any moral standing whatsoever. Some people might intuit, to the contrary, that all human beings have moral standing, including anencephalic infants. However, I think that this intuition is considerably less (a) widespread, and (b) robust than the intuition that severely impaired human beings with similar cognitive capacities to the lower animals have a moral status equal or comparable to that of typical human beings.

An anonymous reviewer asks whether a unitarian (like myself) who claims that human beings typically have lives far richer in well-being than those of the lower animals is thereby forced to accept a view almost as implausible as the first horn of the anti-unitarian's dilemma, viz. the view that one should usually prioritise preventing typical human beings from suffering physical harms (like toothaches, or getting killed) over preventing cognitively impaired human beings from suffering these physical harms, since typical human beings will usually have lives richer in well-being than the lives of impaired human beings.

In reply, let me say first of all that I do not think that the following claim: '(a) one should typically give unimpaired human beings priority over impaired human beings in virtue of the fact that they will typically have stronger interests at stake' is anywhere near as implausible as this claim: '(b) the interests of an unimpaired human being matter more than the interests of an impaired human being, even when the impaired human being's interests are just as strong,' (claim (b) follows from axiological hierarchy plus the first horn of the anti-unitarian's dilemma). On the contrary, I think that many people who are attracted to consequentialism will regard claim (a) as intuitively plausible.

A non-consequentialist unitarian, on the other hand, can avoid claim (a) if she wants to, even if she believes that unimpaired human beings will typically have lives richer in well-being

than those of impaired human beings. She can claim, for instance, that we should prioritise averting R's toothache over averting S's toothache only if the difference in welfare at stake is u utils or greater. Selecting a suitable value of u will imply that we should prioritise averting an unimpaired human being's toothache over averting a cockroach's toothache, but not over averting an impaired human being's toothache. Of course, if the welfare loss caused by an impaired human being's toothache is equal to the welfare loss caused by a dog's toothache, then the non-consequentialist unitarian will have to say that we should not prioritise averting an impaired or unimpaired human being's toothache over averting a dog's toothache. But I think that most non-consequentialist unitarians will claim that this implication is not nearly as implausible as the first horn of the anti-unitarian's dilemma.

VIII. Conclusion

In this paper, I have drawn attention to one respect (success in avoiding the anti-unitarian's dilemma) in which unitarianism appears to be superior to anti-unitarianism. Of course, I cannot hope to have come close here to settling the debate between hierarchy and unitarianism, since I lack the space to discuss several of the arguments relevant to that debate. Kagan himself, for example, has advanced an important 'argument from distribution' in favour of hierarchy (HCA, chapter 3), that I lack the space to discuss in this paper.[42] Nonetheless, I hope that the arguments that I have advanced in this paper can form part of a broader defence of unitarianism, in the face of sophisticated new hierarchy theories like Kagan's.

---

[1] Shelly Kagan, *How to Count Animals, more or less* (Oxford: Oxford University Press, 2019). In-text citations with the label 'HCA' will be to this volume.

[2] Shelly Kagan, 'What's wrong with speciesism? (Society for Applied Philosophy Annual Lecture 2015)', *Journal of Applied Philosophy*, 33, 1 (2016): 1-21.

[3] Another aspect of Kagan's axiological hierarchy that I won't discuss in this paper is his view that our principles of *distributive justice* should differentiate between human beings and lower animals. I hope to critique this form of hierarchy in a future paper (#1), tentatively titled 'Animals, distributive justice, misfortune, and the nature of inequality'. See also Huub Brouwer and Willem van der Deijl, 'All animals are equal, but some more than others?', *Journal of Moral Philosophy*, 17, 3 (2020): §3.

[4] For a detailed discussion of several possible forms of rights-focused hierarchy, see HCA, chapter 9.

[5] Immanuel Kant, *Lectures on Ethics*, trans. P. Heath (Cambridge: Cambridge University Press, 1997), pp. 212-3.

[6] Peter Singer is perhaps the most famous contemporary advocate for unitarianism.

[7] Jean Kazez, *Animalkind: What We Owe to Animals* (Oxford: Wiley-Blackwell, 2010), pp. 91-3.

[8] This is certainly not the only argument that Kagan adduces in favour of axiological hierarchy. Most importantly, Kagan also appeals to an *argument from distribution* [HCA, p. 98-9], which I won't discuss in this paper, since I lack the space to give it the critical attention that it deserves. Cf., however, Lloyd manuscript op. cit., and Brouwer and van der Deijl op. cit..

[9] This is a widely acknowledged point – for a list of references, see HCA, p. 48.

[10] As Jeff McMahan has put it, pain typically has a higher 'opportunity cost' for human beings than it does for animals – see Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (Oxford: Oxford University Press, 2002), p. 229. In-text citations with the label 'EK' will be to this volume.

[11] In *How to Count Animals*, Kagan reports that he "finds it impossible to construct a case" about which we can be one-hundred-percent confident "that the person and the animal have the same amount of well-being at stake" (HCA, pp. 100-1). In proposing my toe-stubbing case, I hope to meet this challenge: making the pains in question short in duration seems to me to be the crux.

[12] When I first presented this thought experiment to Shelly Kagan (in conversation), he reported to me that he shared this intuition. In more recent correspondence, however, Kagan reports to me that when he "think[s] about the case now," he finds himself "inclined to think that the human pain does count more, though … this is a less powerful intuition" than his intuition in the toothache case.

[13] I thank Shelly Kagan for pressing me to clarify this point.

[14] Karsten Stueber, 'Empathy' in Edward N. Zalta (ed.) *Stanford Encyclopedia of Philosophy* (Fall 2019); Abigail A. Marsh, 'The neuroscience of empathy', *Current Opinion in Behavioral Sciences*, 19 (2018): 110-5.

[15] Christine Korsgaard, *Fellow Creatures: Our Obligations to the Other Animals* (Oxford: Oxford University Press, 2018), pp. 13-4, 60-61. For an introduction to some of the evidence that our empathetic faculty is susceptible to irrational biases, see Karsten op. cit., §5.2.

[16] Elisa Silvia Colombo et al., 'Empathy toward animals and belief in animal-human-continuity in Italian veterinary students', *Animal Welfare*, 25, 2 (2016): 275-86.

[17] Bryan W. Van Norden (trans.), *Mengzi: With Selections from Traditional Commentaries* (Indianapolis, IN: Hackett, 2008), p. 9 (my interpolation).

[18] Lucius Caviola et al., 'Utilitarianism for animals, Kantianism for people? Harming animals and humans for the greater good', *Journal of Experimental Psychology: General*, forthcoming.

[19] Here's an example of how we might realise the idea that natural rights can have differing degrees of 'robustness' (cf. also §I above): if saving *k* or more human lives is what it takes to justify harvesting the organs of one unconsenting human being, then saving *k/2* or more dog lives might be all that it takes to justify harvesting the organs of one unconsenting dog.

[20] Although cf. McMahan, who regards sacrificing this intuition as a price worth paying (EK, p. 260) for a theory of rights-focused hierarchy according to which an agent's moral status – in particular whether or not she has a right against being killed – depends upon her actual cognitive capacities (EK, pp. 240-65). McMahan is attracted to the view that typical human beings have a robust right against being killed on account of his intuition that murdering a human being who would otherwise have enjoyed an extremely happy life is no worse, morally speaking, than murdering a human being who would otherwise have enjoyed only a moderately happy life (EK, pp. 233-43). He is attracted to the view that possessing this right requires a certain degree of cognitive development (shared by most human beings but not by most lower animals) on account of his intuition that it would be implausible to afford a similarly-robust natural right against being killed to the lower animals (EK, pp. 245-6).

I think that most people would find it implausible to suppose that severely cognitively impaired human beings should have the kind of lower moral status assigned to them by McMahan's theory. For a critical discussion of McMahan's account of the scope of (certain) natural rights, see Tim Mulgan, 'Critical notice of Jeff McMahan *The Ethics of Killing: Problems at the Margins of Life*', *Canadian Journal of Philosophy*, 34, 3 (2004): 443-60; McMahan concedes the force of Mulgan's line of argument in Jeff McMahan, 'Challenges to human equality', *Journal of Ethics*, 12, 1 (2008): 81-104. See also Richard J. Arneson, 'What, if anything, renders all humans morally equal?' in D. Jamieson (ed.) *Singer and His Critics* (Oxford: Blackwell, 1999), pp. 103-28; and HCA, chapter 8. I hope to argue in a future paper (#2), tentatively titled 'Killing: how consequentialists should respond to the equal wrongness thesis', that non-hierarchical consequentialism can explain and vindicate several of the most plausible 'equal wrongness' intuitions that motivate McMahan's theory without invoking the idea of natural rights.

[21] An anonymous reviewer asks whether by drawing upon Kagan's discussion of 'practical realism' (HCA, §11.2), an actual-capacities theorist might be able to defend the view that one should treat a severely cognitively impaired human being *as if* she has the same moral status as a typical human being. Roughly speaking, *practical realism* is the claim that hierarchy theorists who recognise a continuum of possible moral statuses should nonetheless carry out their day-to-day moral reasoning using a rough and ready model of moral status, according to which there is a low and finite number of moral-status *categories*. Differences in moral status between different unimpaired human beings can usually be safely ignored in our day-to-day moral reasoning, Kagan counsels, because they are likely to be negligibly small.

The difference in capacities between a typical human being and a severely cognitively impaired human being, however, is *not* negligible. A severely cognitively impaired human being has, *ex hypothesi*, cognitive capacities that are similar to those of a lower animal like a cow or a dog. Thus, an actual-capacities practical realist who wants to argue that we should treat severely impaired human beings as though they have the same moral status as typical human beings (even once we are aware of their impairments) will have to maintain that we should also treat cows and dogs as though they have this kind of high moral status. This result is almost certainly contrary to the kind of intuitions that motivate hierarchy in the first place. In any case, it seems hard to see how practical realism could prompt us to adopt a moral-status category that ranges widely enough to include both Einstein and my dog.

[22] For a philosopher's list of twenty-six potential definitions, see John S. Wilkins, *Defining Species: A Sourcebook from Antiquity to Today* (New York: Peter Lang, 2009).

[23] Kagan 2016 op. cit., §3.2.

[24] Ernst Mayr, *Systematics and the Origin of Species: From the Viewpoint of a Zoologist* (New York: Columbia University Press, 1942).

[25] For a much more detailed critique of speciesism, sensitive to several alternative definitions of species, see EK, pp. 211-28.

[26] Christopher J. G. Meacham, 'Person affecting views and saturating counterpart relations', *Philosophical Studies*, 158, 2 (2012): 257-87; Rachel Cooper, 'How might I have been?', *Metaphilosophy*, 46, 4-5 (2015): 495-514; Shamik Dasgupta, 'Essentialism and the nonidentity problem', *Philosophy and Phenomenological Research*, 96, 3 (2016): 540-70. Cf. also David Wasserman, 'Personal identity and the moral appraisal of prenatal therapy' in L. S. Parker and R. A. Ankeny (eds.) *Mutating Concepts, Evolving Disciplines: Genetics, Medicine and Society* (Dordrecht: Kluwer, 2002), pp. 235-64; Anthony Wrigley, 'Genetic selection and modal harms', *Monist*, 89, 4 (2006): 505-25; 'Harm to future persons: non-identity problems and counterpart solutions', *Ethical Theory and Moral Practice*, 15, 2 (2012): 175-90; Nicola Jane Williams, 'Possible persons and the problem of prenatal harm', *Journal of Ethics*, 17, 4 (2013): 355-85; Janet Malek, 'The non-identity non-problem: constructing continuity of identity

among possible persons', *Law, Ethics and Philosophy*, 7 (2019): 15-28. For an introduction to the broader literature on essential versus accidental properties, see Teresa Robertson and Philip Atkins, 'Essential vs. accidental properties' in Edward N. Zalta (ed.) *Stanford Encyclopedia of Philosophy* (Spring 2018).

[27] Those who doubt whether this genetic therapy could be identity-preserving for Fido might find it helpful to suppose that the therapy takes place over a period of several years, so that the improvement in Fido's cognitive capacities that occurs on any given day or week is relatively small. Given the transitivity of identity, this sort of therapy looks likely to be identity-preserving.

[28] For a summary of the critical literature on potential-capacities theories of moral status, see Agnieszka Jaworska and Julie Tannenbaum, 'The grounds of moral status' in Edward N. Zalta (ed.) *Stanford Encyclopedia of Philosophy* (Spring 2018): §5.2. For a critique of Kagan's appeal to potential capacities, see Jeff McMahan, 'On 'modal personism'', *Journal of Applied Philosophy*, 33, 1 (2016): 26-30.

[29] I outlined an argument for this view in the context of rights-focused hierarchy in §III of this paper.

[30] As David DeGrazia has pointed out, it is one thing to claim that the number of field goals scored by one's team should affect the final score, but quite another thing to claim that the number of field goals that one's team *could* have scored should affect the final score – see David DeGrazia, 'Modal personhood and moral status: a reply to Kagan's proposal', *Journal of Applied Philosophy*, 33, 1 (2016): 24. Kagan himself admits that he "would be considerably more comfortable if [he] could go beyond the mere appeal to brute intuition, eventually offering an account of *why* modal personhood should matter in the ways we may intuitively think that it does" (2016 op. cit., §3.8). Adam Roberts emphasises the urgency of this question in Adam James Roberts, 'Pessimism about motivating modal personism', *Journal of Applied Philosophy*, 35, 3 (2018).

[31] This line of argument is suggested by Doran Smolkin in 'Kagan on speciesism and modal personism', *Journal of Applied Philosophy*, 36, 1 (2019): 83, 86; and 'Is humane farming morally permissible?', unpublished manuscript. See also Frederik Kaufman, 'Speciesism and the argument from misfortune', *Journal of Applied Philosophy*, 15, 2 (1998).

[32] Smolkin (2019 op. cit.) interprets Kagan as endorsing this view in Shelly Kagan, 'Shelly Kagan on death and deprivation', *Philosophy Bites*, podcast, April 2017; cf. Kai Draper 'Disappointment, sadness, and death', *Philosophical Review* 108, 3 (1999): §2.

[33] I thank Doran Smolkin for suggesting this example.

[34] I am suggesting here that what is sometimes referred to as the *lack of control conception* of luck is the reading of that concept that is relevant to moral status and to egalitarian justice. For a useful introduction to the philosophical literature on luck, see Fernando Broncano-Berrocal, 'Luck' in *The Internet Encyclopedia of Philosophy*, https://www.iep.utm.edu/, accessed 20th April 2020.

[35] I defend this apparently absurd view in Lloyd manuscript #1 op. cit..

[36] Jeff McMahan defends this view, although not by defending the modal account of misfortune (see below) that would be helpful to the modal capacities hierarchy theorist (EK, pp. 145-65, 312-4). A hierarchy theorist who concedes the force of my arguments against Kagan's modal capacities theory might wonder whether she could develop an alternative theory of hierarchy, according to which an impaired human being deserves high moral status because her limited cognitive endowment presages a misfortune according to McMahan's *intrinsic potential account* of misfortune (to be clear: McMahan himself does not endorse this kind of theory). Unfortunately for the hierarchy theorist, however, any version of this new theory that puts significant moral distance between typical human beings and the lower animals is going to be impaled upon the first horn of the anti-unitarian's dilemma, since McMahan's intrinsic

potential account of misfortune implies that a cognitively impaired human being is unfortunate only if her impairment is non-congenital (EK, p. 153). I critique this aspect of McMahan's theory of misfortune in Lloyd manuscript #1 op. cit..

[37] Kagan himself has described the "asymmetry" in his theory of moral status as "puzzling" (HCA, p. 141). The fact that the asymmetry *remains* puzzling even for somebody who adopts the luck-based defence of Kagan's theory suggests, however, that this defence is at best incomplete; a better defence of Kagan's theory would surely afford us some kind of satisfying explanation for the theory's 'puzzling' asymmetricity.

[38] I borrow this example from Smolkin 2019 op. cit., p. 86.

[39] Cf. Daniel Kahneman and Amos Tversky, 'The simulation heuristic' in D. Kahneman, P. Slovic and A. Tversky (eds.) *Judgement Under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press, 1982), p. 203.

[40] It is important to recall here that I am using 'fortune' as a technical term, to refer to how well an individual's life is going (EK, p. 145). Missing out on some great good is, at least in the everyday sense of the word, of course *unluckier* for someone who stood a strong chance of achieving that good than for someone who stood no chance at all. Indeed, this is precisely why the former sort of person is likely to be more annoyed and frustrated at missing out on the good in question than the later sort of person would be.

[41] For critical discussions of two of the most promising hierarchical theories of moral status to have been proposed prior to *How to Count Animals*, see Christopher Grau, 'Moral status, speciesism, and Liao's genetic account', *Journal of Moral Philosophy* 7, 3 (2010): 387-96; David DeGrazia, 'On the moral status of infants and the cognitively disabled: a reply to Jaworska and Tannenbaum', *Ethics* 124, 3 (2014): 543-56; Travis Timmerman and Bob Fischer, 'The problem with person-rearing accounts of moral status', *Thought* 8, 2 (2019): 119-28.

[42] But cf. Lloyd manuscript #1 op. cit., and Brouwer and van der Deijl op. cit..