

PROBLEMS WITH "PERSONS"

by

PHILLIP COLE*

1. *Introduction*

This paper critically examines the concept of a person that has come to play a central role in medical ethics. The dominant view is that the most valuable lives belong to what can be termed "persons"; that personhood amounts to the possession of certain capacities that only persons possess; and that those capacities are rationality and self-consciousness or self-awareness. This view also holds that only persons are morally wronged by being killed, because only persons have an interest in continued life by virtue of their self-awareness. Anything that lacks rational self-consciousness is a non-person, and cannot be morally wronged by being killed, although they can be wronged in other ways. Such a being cannot have an interest in continued existence, but can have other interests (for example not suffering unnecessary pain or distress) and frustrating these interests can count as a moral wrong. It is important to note, therefore, that this dominant version of "personhood" does not claim that it is impossible to do moral wrongs to non-persons, and that therefore that we can do what we like to them — the claim is only that non-persons are not morally harmed when they are killed.

Two key figures in developing this position are John Harris¹ and

* Centre for Practical Philosophy, School of Philosophy and Religious Studies, Middlesex University. The author would like to thank *Res Publica's* anonymous readers for reading and most helpfully commenting on earlier drafts. The paper develops ideas first expressed in Phillip Cole, "Dangerous Philosophies", *Radical Philosophy* 72 (1995), July/August, 44–46; and in "The Futility of Bioethics", a paper presented at the Consensus in Bioethics conference, University of Central Lancashire (October 1996). However, it constitutes a substantial revision of those earlier ideas.

1 J. Harris, *The Value of Life: an Introduction to Medical Ethics* (London: Routledge and Kegan Paul, 1985).

Peter Singer.² Harris says:

[T]o value its own life, a being would have to be aware of itself as an independent centre of consciousness, existing over time with a future that it was capable of envisaging and wishing to experience. Only if it could envisage the future could a being want life to go on, and so value its continued existence.³

Singer says: "... I propose to use 'person', in the sense of a rational self-conscious being";⁴ and: "[K]illing a person who prefers to continue living is ... wrong, other things being equal."⁵ And "[I]n contrast, beings who cannot see themselves as entities with a future cannot have any preferences about their own future existence."⁶

The purpose of this paper is not to reject the concept of a person, or to reject the view that rationality and self-consciousness are relevant to personhood. My main objective, rather, is to dispute the claim that the concept of a person can play a particular role in medical ethics, that is, that it can settle the morality of killing. The specific claim I want to dispute is that while killing persons is, on balance, to do them a moral wrong, killing non-persons cannot, on balance, do them a moral wrong, because only the former have an interest in continued existence. The dominant view places a great deal of weight on the coincidence between the persons/non-persons boundary and the immorality/acceptability of killing boundary: my argument is that these boundaries do not coincide. In arguing this, however, I do want to make certain points about the *nature* of the concept of personhood, and suggest that a great deal rests on exactly how it is understood. In Section 2, I examine John Harris's arguments about the nature of personhood in some depth. In Section 3, I argue against the view that the concept carries the moral force that the dominant position assumes it carries, looking in particular at the policy arguments of Dan Brock.⁷ In Section 4, I conclude by sketching out what I take to be a plausible conception of the person, and make some observations about the moral force of *that* conception.

2 P. Singer, *Practical Ethics* (Cambridge: Cambridge University Press, 1993, 2nd ed.).

3 *Supra* n.1, at 18.

4 *Supra* n.2, at 87.

5 *Supra* n.2, at 94.

6 *Supra* n.2, at 95.

7 D.W. Brock, *Life and Death: Philosophical Essays in Biomedical Ethics* (Cambridge: Cambridge University Press, 1993).

2. *John Harris's Approach*

John Harris appeals to two methods of establishing what makes life valuable. The first rests on looking at persons and establishing what makes them morally distinct from non-persons, and Harris claims that this method shows that a combination of rationality and self-consciousness constitutes the essential element of a valuable life. The second method asks what makes existence valuable for persons, and Harris concludes that, while this question has no answer, asking it nevertheless leads to the conclusion that persons are beings that have the capacity to value their own existence. It is the "ability to want to experience the future, or to want not to experience it and the awareness of those wants"⁸ that is important — killing such a being has the potential to frustrate those wants, and therefore morally wrongs them. "Creatures that cannot value their own existence cannot be wronged in this way, for their death deprives them of nothing that they can value."⁹ Harris argues that the two strategies converge, because when we ask what it takes to be the sort of being that can value its own existence, it turns out to be the possession of the properties of rationality and self-consciousness.¹⁰ In this section I examine Harris's first strategy, in order to explore the nature of the concept of a person.

Harris's first strategy "involves looking at what it is that's so different about a person that justifies our valuing such a creature above others".¹¹ This method "concentrates on examining which of the differences between people and other creatures seem relevant to the question of their differential value".¹² And:

[I]f we look at creatures we are sure are persons if any are — normal adult human beings — and can find features of their lives or capacities which ... incline us to judge their lives of more significance and value than lives which lack such features, we might come close to a concept of the person.¹³

Harris goes on to describe this strategy as "identifying the defining

8 *Supra* n.1, at 18.

9 *Supra* n.1, at 19.

10 *Supra* n.1, at 18.

11 *Supra* n.1, at 14.

12 *Supra* n.1, at 14.

13 *Supra* n.1, at 14.

characteristics of a person”¹⁴ and he cites John Locke as providing a way to “distinguish persons from other creatures in a way that made sense of the difference in value we place upon them”.¹⁵ Locke makes “a combination of rationality and self-consciousness the distinguishing features of a person” and Harris says this has “much to recommend it”.¹⁶ What he is looking for is “an account of what it is to be a person” that manages to “isolate features that were ... valued by us and which were of the sort to account for the peculiar status that we give to creatures possessing such features”.¹⁷

Now, there are some puzzles with this strategy. Harris begins by claiming that it is a strategy for deciding what it is that gives persons their special moral value, and that what the strategy tells us is that it is rational self-consciousness that gives persons their special moral value. However, it can also be interpreted as a strategy for defining persons, for distinguishing between persons and non-persons — and again the answer is that the defining property is rational self-consciousness. There are potentially, therefore, two questions here: (1) What makes persons morally valuable?; and (2) What makes something a person? These are different questions, and we cannot assume that what counts as an answer to (1) will also count as an answer to (2).

What happens if we interpret Harris as trying to supply an answer to the second question, trying, that is, to “distinguish persons from other creatures”? One way of understanding the task is to arrive at a set of necessary and sufficient conditions for being a person, and Harris could be saying that rationality and self-consciousness are such conditions. But will his method get us to that conclusion? Interpreted this way, the argument would have the following form:

1. We are trying to define Xs (persons).
2. We look for instances of unproblematic Xs, and note those properties which lead us to want to judge them to be Xs.
3. We note that we want to judge them to be Xs, because they possess property A (rational self-consciousness).
4. Therefore we conclude that A is important to X-hood, such that

14 *Supra* n.1, at 14–15.

15 *Supra* n.1, at 15.

16 *Supra* n.1, at 15.

17 *Supra* n.1, at 15.

anything that possesses A can be judged to be X.

5. We also conclude that A is important to X-hood, such that anything that does not possess it cannot be judged to be X.
6. Therefore, under (4), only Xs must have property A so that A is a *sufficient* condition of X-hood.
7. And, under (5), all Xs must have property A and therefore A is a *necessary* condition of X-hood.

The problem with the argument in this form is that the move from (3) to (4) and (5) is not logical, but intuitive, which is why (6) asserts what is not the logical case. Harris (under this interpretation) is looking for properties that "incline us to judge" something to be a person. I am inclined to judge that this subject is an X because it possesses property A. On the strength of that inclination, can I then claim that all Xs have property A, and that only Xs have property A? How far I get surely depends upon the strength of the inclination. For example, what inclines me to judge that this is a philosophy paper by Professor Smith is that the name "Professor Smith" is written on the cover sheet. On the strength of that inclination, can I then claim that having "Professor Smith" on the cover sheet is a sufficient condition to show that it is a paper by Professor Smith — that *only* papers by Professor Smith have "Professor Smith" on their cover sheet? Even if so, this still allows that some papers by Professor Smith do not have "Professor Smith" on the cover sheet (call this alternative I). Can I also claim that having "Professor Smith" on the cover sheet is a necessary condition to show that it is a paper by Professor Smith — that all papers by Professor Smith have "Professor Smith" on their cover sheet? If so, this allows that papers other than papers by Professor Smith have "Professor Smith" on their cover sheet (call this alternative II). The point is that if alternative I is true, then this rules out the property as a necessary condition, and if alternative II is true, this rules it out as a sufficient condition; and moreover, both alternatives seem reasonably plausible.

On the evidence so far, it remains plausible that some of Professor Smith's papers do not have her name on their cover sheet (alternative I), and therefore the property cannot act as a necessary condition. And it also seems plausible that some papers could have her name on the cover sheet and not be by her (alternative II) — it may be that it is a student paper for Professor Smith's attention, where the student forgot to put their own name on the cover sheet as well (these things happen!). Thus

the property cannot act as a sufficient condition either. Therefore the fact that property A leads me to judge that this is an X in this case does not show that the property is either a necessary or sufficient condition for being an X. Therefore the move from (3) to (4) and (5) in the version of Harris's argument presented above is neither logically *nor* intuitively compelling.

What if we accept the first interpretation of what Harris is doing — supplying an answer to the question, “What is it that leads us to judge that persons are morally valuable?” Again the answer is that persons are rationally self-conscious. Can we interpret this as an argument to show that rationality and self-consciousness are necessary and sufficient conditions for the moral value of persons? The argument under this interpretation would have the following form:

1. What inclines us to judge that Xs (persons) are Y (morally valuable)?
2. We look for instances of unproblematic Xs and note those properties that lead us to want to judge them to be Y.
3. We note that we want to judge Xs to be Y because they possess property A (rational self-consciousness).
4. Therefore anything that possesses A must be Y.
5. Therefore anything that does not possess A cannot be judged to be Y.
6. Therefore, under (4), only things that possess A are Y, and A is a sufficient condition for Y (rational self-consciousness is a sufficient condition for moral value).
7. And, under (5), all things that are Y must possess A, and A is a necessary condition for Y (rational self-consciousness is a necessary condition for moral value).

Again the move from (3) to (4) and (5) is not logical, but intuitive; again (6) asserts what is not logically so; and again the initial move must rest upon the strength of the intuition. The fact that persons possess rationality and self-consciousness leads me to judge that persons have moral value — but does this compel me to the judgment that *anything* that possesses rational self-consciousness must have moral value, that it is a sufficient condition? Does it compel me to the judgment that *only* things that possess rational self-consciousness have moral value, that it is a necessary condition? While it logically compels me to neither, the

move from (3) to (4), the claim that rational self-consciousness is a *sufficient* condition of moral value, looks a reasonably strong intuition. It is still not *logically* compelling, however, because it remains possible that a property that makes X morally valuable would not make something other than X morally valuable. The argument is not "if A then Y" — it is "if X plus A, then Y", and we cannot conclude from that, that "if Z plus A, then Y". For example, the properties that make a philosophical paper valuable may not make other kinds of papers valuable, and vice versa. Therefore the move from (3) to (4) remains logically unconvincing, although I do find it intuitively a stronger move than the equivalent stage in the alternative version of the argument. It would mean that anything that was rationally self-conscious would be morally valuable: (i) this rules out the possibility of anything being rationally self-conscious and not being morally valuable; and (ii) it allows that something could lack rational self-consciousness and yet be morally valuable. Both (i) and (ii) strike me as plausible.

However, I do not find the move to (5), the view that rational self-consciousness is a *necessary* condition of moral value, intuitively compelling at all. The fact that persons possess rationality and self-consciousness leads me to judge that persons have moral value, but this doesn't move me towards the judgment that in that case *only* things that possess rationality and self-consciousness have moral value. It would mean that in order to be morally valuable something would have to be rationally self-conscious: (i) this simply *rules out* something being morally valuable if it is not rationally self-conscious; and (ii) it allows that some things that are rationally self-conscious have no moral value. Both (i) and (ii) strike me as extremely implausible.

How damaging is all this to Harris's account of a person? Strictly speaking, neither of the above interpretations can be imposed upon his first strategy, but it is important to *note* that they can't. Furthermore, although Harris does not talk in terms of necessary and sufficient conditions for personhood or of the moral value of personhood, he does say of his strategy that it identifies the "defining characteristics of a person";¹⁸ and that the combination of rationality and self-consciousness are "the distinguishing features" of persons.¹⁹ At its most modest, all

18 *Supra* n.1, at 14–15.

19 *Supra* n.1, at 15. Note that Singer also does not explicitly claim that rationality and self-consciousness are necessary and sufficient conditions for personhood, only that he accepts "definitions" in these terms, and that "I

Harris's strategy suggests is that one of the features that marks persons as morally different from non-persons is their possession of rational self-consciousness. But three much less modest claims might be thought to follow from this: (1) that the *only* property that makes persons morally different from non-persons is rational self-consciousness; (2) that rational self-consciousness is a necessary and sufficient condition of the moral value of persons; and (3) that rational self-consciousness is a necessary and sufficient condition of personhood. What I have tried to show is that none of these three claims can emerge from Harris's strategy, and that none of them are anyway remotely plausible. Of the three claims, Harris himself may be making the first, but I am not suggesting that he makes either of the others.

An additional problem with the strategy, and perhaps the problem that underlies the above concerns, is that it presupposes that we have already decided the defining characteristics of personhood — otherwise how could we have made the distinction between persons and non-persons which both versions of Harris's strategy begin with? Harris chooses "normal adult human beings" as indisputable persons — but why? Could it be that his choice has something to do with the fact that they are rationally self-conscious? But it is not as though we decide that normal adult human beings are indisputable persons, and then examine them and discover that they happen to possess rational self-consciousness: we cannot *discover* that rational self-consciousness is the essence of personhood; we can only *decide* that it is. We can *construct* the concept of personhood around a certain set of features, but we cannot *discover* that these features are the essence of personhood — we can only discover what *makes us want to call* certain beings persons. The notion of a person is not one that is discovered, but one that is recognised as such.

In his discussion Harris draws on the approach of Clifford Grobstein, who does talk of the importance of recognition by others:

[T]he question is to determine when in development the embryo or foetus is generally recognizable as human and evokes empathy as another self The question thus becomes the stage at which the embryo or foetus can

propose to use 'person', in the sense of a rational and self-conscious being" — *supra* n.2, at 87.

first be generally recognized as human, generating empathy as a person or a person to be.²⁰

In other words, the issue is what makes us want to call something a person? What makes us want to offer recognition? Harris is unhappy with this aspect of Grobstein's approach:

he has no reason to suppose that recognising a foetus (or anything) as distinctly humanoid is any evidence at all of the presence of the requisite inner qualities.²¹

In other words, the foetus, or anything, has to earn that recognition by possessing specific properties which we must discover by empirical observation. Harris continues:

[W]hy should we not assume that the empathy evoked by the sight of the three-month-old foetus is just the soggy sentimentality classically evoked by proximity to dependent sentient creatures, like puppies?²²

And he concludes:

[I]f the social and moral status of person is "accorded through recognition and acceptance by others", it is so accorded we may hope with justification. People are not people because they are accepted, but rather they are accepted because they are people.²³

Harris is surely right that in advance of the recognition of something as a person, we must already have some idea of the features that constitute personhood. But it does not follow from this that the features which constitute personhood can be discovered by investigation, because personhood is not an empirical fact about certain sorts of creatures — it is a moral construct. We construct the concept of personhood around a set of features, one of which may well be rational self-consciousness, but that construction is a moral process of creation, not an empirical process of discovery. In the end, we have to come to understand why rational self-consciousness makes us want to call such beings "persons", and we have to decide whether the absence of that property carries the moral significance claimed for it.

To conclude this section, "personhood" is a moral concept, and not an empirical one. We do not discover persons, we create them; and that

20 C. Grobstein, *From Chance to Purpose* (London: Addison Wesley, 1981), 102.

21 *Supra* n.1, at 13.

22 *Supra* n.1, at 13.

23 *Supra* n.1, at 14.

act of creation is constituted by the intuitive recognition of a pattern of features. The concept of a person is a concept of recognition, and what we recognize is that there is some cluster of features here arranged in some way that we can identify with — and while this process is vague, complex, and perhaps unphilosophical, it may be the best that we can do. A clear definition of a person in terms of necessary and sufficient conditions is of limited use here, and the empirical discovery of whether a particular being possesses these conditions is an only partially accurate description of what we are doing when we identify something as a person. Even though we may, in such an individual case, be *discovering* whether that being possesses a set of features, we would have to keep in mind that an essential part of what we are doing is an intuitive recognition of it as a person, rather than a process of ticking off features against a checklist. In a sense we have to “learn to see” persons — nobody can supply us with a definitive checklist of features and then send us out into the world to discover persons.

On the basis of this understanding of the nature of the concept, I have no argument with the notion of a person; no argument with the claim that personhood brings with it a special moral status; and indeed no argument with the claim that rationality and self-consciousness have an important role to play here. My argument is with an understanding of the concept that sees it as consisting of a definitive set of necessary and sufficient conditions, and which claims that these conditions can be discovered by empirical investigation of particular cases, which can then be applied to particular cases to discover persons empirically. Instead, it is a cluster of features that we intuitively take to be important in certain combinations; and we can recognize, partly by creative intuition and partly by empirical discovery, that in particular cases those features are present in a significant pattern. My remaining argument is with the claim that the boundary between persons/non-persons coincides with the boundary between the immorality/acceptability of killing.

3. *Dan Brock's Application*

I want to approach that argument by examining Dan Brock's application of the concept of a person to a particular problem in medical ethics. One of his concerns is whether severely demented people have any just claim to the health care resources needed to sustain life. He answers this question by making a series of comparisons. He begins with

the uncontroversial case of dead people, who obviously have no interest in continuing life, and therefore no just claim to the necessary resources. Brock then considers people in persistent vegetative states, and claims that

some degree of present or future capacity for sentience, conscious experience, or purposive behaviour is uncontroversially necessary for being a person in the sense relevant to being an independent source of claims within theories of distributive justice.²⁴

Persistent Vegetative State (PVS) patients have lost any such capacity:

[S]o far as their present or future capacities for any psychological life or conscious experience are concerned, PVS patients are similar to such living things as plants, which lack any capacity for sentience or a conscious life; they lack the sentience that even many animals have.²⁵

Therefore they cannot possibly have any interest in claims to health care to prolong life.

The severely demented are defined by Brock as those whose cognitive capacities and memories are so destroyed as to eliminate any psychological continuity. Since Brock uses a theory of personal identity based on psychological continuity, it follows that while these people were persons in the past, they cannot be persons now, because there is no psychological continuity here — they cannot make links across time, and they therefore lack personhood. They “approach more closely the condition of animals than normal adult humans in their psychological capacities”.²⁶ Animals are not wronged when they are killed, because they lack the capacity to experience themselves as a single self-conscious individual who persists through time. The severely demented also lack this capacity, and so they, too, are not morally wronged if they are killed: “they lose the fundamental basis for persons’ interest in continued life and in measures that sustain life — that their future life is a necessary condition for satisfying all of a person’s desires about and plans for the future”.²⁷

Brock emphasises that this is not a judgment about the quality of life such people can expect. Rather it is to say that such people are not persons: therefore “the severely demented have lost an interest in

24. *Supra* n.7, at 368.

25. *Supra* n.7, at 364.

26. *Supra* n.7, at 372–73.

27. *Supra* n.7, at 373.

treatment whose ultimate purpose is to prolong or sustain their lives".²⁸ This is not to say that they have no rights to any medical care — they have an interest in measures that keep them clean, for example, and in any measure "that treats them with dignity out of respect for the person they once were".²⁹

However, this last claim opens a serious gap in Brock's approach. He is claiming that only rational self-conscious beings have valuable lives of the kind that merit respect in terms of not being killed. However, he has allowed that it is not only the current or future possession of rational self-consciousness that can give rise to respect. The former possession of it can carry significance too; the severely demented are entitled to dignity "out of respect for the persons they once were". This also applies to the dead. One cannot assume that the dead have no rights; indeed much legislation is based on the assumption that they do. Brock observes that

it is generally acknowledged that the bodies even of the dead must be treated with dignity respecting the person, and the person's wishes, whose body this once was.³⁰

Not only are the dead entitled to respect, but one aspect of that respect will be respect for that person's wishes. But if the dead are entitled to respect on this basis, then so are PVS patients and the severely demented. And if this is correct, then one cannot base a decision to end life-sustaining treatment only on a view of their current state or potential future state — i.e., that they are not and never will be a person. One must also show respect for them on the basis of their past possession of personhood, and part of that respect will be for any past wishes they may have expressed.

This means that if a person expressly wished that they should not be killed or have life-sustaining treatment withdrawn if they were to become severely demented or enter a persistent vegetative state, then those wishes must, to some extent, be respected. They need not, perhaps, determine the decision, but they must enter into it, inasmuch as they do supply a morally good reason not to kill that person or end their treatment. That reason may be outweighed by other factors, but it remains nevertheless a morally good reason, such that killing them or withdrawing their treatment in the absence of any moral reasons which

28 *Supra* n.7, at 375.

29 *Supra* n.7, at 375.

30 *Supra* n.7, at 376.

outweigh it does them a moral wrong. That means it is perfectly possible that ending the life of a being that is not rationally self-conscious and never will be again does them a moral wrong — and that the conclusion of the "personhood" argument, that ending the life of a non-person cannot do them a moral wrong, is simply false.

Of course it also follows that if a person expressly wished that they *should* be killed, or *should* have life-sustaining treatment withdrawn if they were to enter a PVS or become severely demented, then this is a morally good reason to do so. Again, respect for their wishes may be outweighed by other moral reasons, but to fail to end their life in the absence of such factors would be to do them a moral wrong. And if a person failed to express a view, then those responsible for their welfare must try to reach a decision that takes into account a view of what they might reasonably be thought to have wanted.

If this is correct, then ending the life of a PVS patient or someone who is severely demented is not as straightforward as Brock suggests: it is simply false to claim that ending the life of such a being cannot do them a moral wrong. Any such decision must be based, not on some calculation of whether they measure up as persons, but upon a range of factors including respect for who they were and respect for what their wishes were or would likely have been. Any decision to end life which does not take such wishes into account is morally wrong.

To conclude this section, I have not shown that there is anything necessarily wrong with Brock's concept of a person; neither have I attempted to show that the severely demented and people in persistent vegetative states *are* persons. What my argument has shown is that there are at least two cases where ending the life of a non-person does them a moral wrong, and therefore that the moral force of the concept of a person has been thrown into doubt. The dominant view holds as a principle that while the killing of a person does them a moral harm, the killing of non-persons can do them no moral harm. This has been shown to be false in at least two instances. This is by no means trivial, as a great deal of weight is placed on the supposed coincidence between the persons/non-persons boundary and the immorality/acceptability of killing boundary. However, it might be thought that the above conclusion is trivial, because I have isolated only two cases that are exceptional, and because even these cases rely upon the former personhood of these particular non-persons. I think, however, that this allegation of triviality would be wrong for two reasons: firstly, the two

exceptional cases are extremely important ones in medical ethics, and the rejection of the coincidence of the two boundaries in these cases has deep implications for practice; and secondly, having shown that particular kinds of non-persons can have an interest in continued existence that makes killing them wrong, we may be able to identify a wider range of non-persons capable of possessing that interest. I shall attempt to make out such a case in the next section.

4. *Conclusion*

I have tried so far to show, firstly, that we cannot arrive at the concept of a person through empirical and/or logical methods, that it is a moral creation, not a logical deduction or an empirical discovery; and secondly, that the concept does not have the moral force it is held to have by some of the major writers in the field. In conclusion, I shall attempt to reinforce the second point, by showing that non-persons can have an interest in continued existence which gives rise to its being morally wrong to kill them. I shall also develop further the “cluster” concept of a person I invoked in section 2, and examine in brief the moral implications of such an approach.

(i) *The immorality/acceptability of killing*

The dominant personhood argument is interest-based, and it identifies one particular interest as determining the wrongness of killing — the interest in continued existence. It attributes that interest only to beings who are rationally self-conscious. Now, it is important to notice that the position cannot hold that beings can have only interests of which they are consciously aware, because these writers would allow that, for example, merely sentient beings have interests in avoiding unnecessary pain, and they need not be consciously aware of this interest in order to hold it.³¹ Singer does say that we should make the plausible move of “taking a person’s interests to be what, on balance and after reflection on all the relevant facts, a person prefers”.³² But he cannot be claiming that interests can be only conscious preferences, because for

31 How could a sentient non-person be consciously aware of an interest in avoiding *unnecessary* pain?

32 *Supra* n.2, at 94.

Singer sentience brings with it interests. It seems implausible to claim that beings can have only interests of which they are consciously aware, and I assume that this is not the argument being put forward here.

Rather, persons have an interest in future existence *because* they are self-consciously aware; and it is not their self-conscious awareness of that interest that matters, but their self-conscious awareness of their possible future existence. The interest arises from *that* awareness, not from an awareness of the interest itself. To claim that they must be aware of the interest itself for it to count as an interest is to add something unnecessarily implausible to the argument. This is important because it now allows that beings can have interests without having to be aware of them. For Singer, in order to have interests one must be at least conscious,³³ but it does not follow from this that one must be conscious *of one's interests*. Singer has three levels of being: non-persons that are non-sentient; non-persons that are sentient; and persons. It is not possible to morally harm the first category at all; it is possible to harm the second category, but not by killing them; and it is only the third category that can be morally harmed by being killed. For Singer, consciousness is necessary because his preference utilitarianism makes no sense without it: but without a commitment to such a moral theory, it does become plausible to attach interests to beings that are not conscious in terms of sentience.

Kenneth E. Goodpaster argues that trees and plants, for example, can be held to have interests:

[T]here is no absurdity in imagining the representation of the needs of a tree for sun and water in the face of a proposal to cut it down or pave its immediate radius for a parking lot. We might of course, on reflection, decide to go ahead and cut it down or do the paving, but there is hardly an intelligibility problem about representing the tree's interest in our deciding not to. In the face of their obvious tendencies to maintain and heal themselves, it is very difficult to reject the idea of interests on the part of trees (and plants generally) in remaining alive.³⁴

And:

the interest principle either grows to fit what we might call a "life principle" or requires an arbitrary stipulation of psychological capacities (for desires,

³³ *Supra* n.2, at 276–80.

³⁴ K.E. Goodpaster, "On Being Morally Considerable", in D. VandeVeer and C. Pierce, eds., *The Environmental Ethics and Policy Book* (Belmont, California: Wadsworth Publishing Company, 1994), 105–12, at 109.

wants, etc.) which are neither warranted ... nor independently plausible.³⁵

One might try to articulate the idea by identifying an interest in becoming a flourishing member of one's kind: this is an interest that does not require consciousness. An interest in continued existence could arise from this, or it could be based on a cluster of interests that non-persons are capable of having. These remarks are, of course, extremely sketchy, but even if one rejects the plausibility of non-sentient beings having interests, the dominant personhood position does allow that sentient non-persons have interests, and those interests could well include a cluster that can plausibly act as a basis for an interest in continued existence. Therefore there would be no justification in limiting the interest in continued existence to beings that are rationally self-conscious.

It might be objected that such an approach would be immensely impractical, in that it would make killing anything that has life morally wrong. However, it does not need to have that implication, for strictly speaking it holds that killing a being with an interest in continued existence is *prima facie* morally wrong — it allows that there could, nonetheless, be good reasons for killing which justify over-riding that interest. No radical "sanctity of life" view is entailed here. For example, a pro-abortion view is perfectly consistent with this position.

What the position does rule out is the claim that killing non-persons, however they are defined, cannot be morally wrong — it can be. It depends on the circumstances and on the reasons and other interests at stake; but it remains perfectly plausible to argue that in particular cases killing a non-person emerges as morally wrong, as it does it a moral harm which is not outweighed by other interests or by any other good reasons. Therefore the immorality/acceptability of killing boundary simply cannot coincide with the persons/non-persons boundary.

(ii) *The "cluster" concept of a person*

In section 2 I suggested that the concept of a person is a "cluster" concept — that it is based on a cluster of features that cannot be reduced to necessary and sufficient conditions. Michael Fox suggests the following for inclusion in such a cluster:

³⁵ *Supra* n.34, at 110.

critical self-awareness; the ability to manipulate complex concepts and to use a sophisticated language ...; and the capacity to reflect, plan, deliberate, choose, and accept responsibility for acting.³⁶

And he suggests that

the possession of these characteristics, plus the capacity to recognize them in others and to care about others, goes a long way toward explaining what we mean by speaking of ourselves as *persons*.³⁷

Another example of this approach is Mary Anne Warren's, who lists sentience, emotionality, reason, the capacity to communicate, self-awareness and moral agency, with wide-ranging interpretations of what these amount to.³⁸ These are in no sense meant as a set of necessary and sufficient conditions: they are rather a cluster which can be recognized in particular patterns, and not all need be present.

This seems to me to be the most plausible approach to personhood. But it has to be pointed out that the claim made by Fox about its moral force — that it determines membership of the moral community — does not follow. Fox defines a moral community as follows:

[M]ost generally, it is a group of beings that shares certain characteristics and whose members are or consider themselves to be bound to observe certain rules of conduct in relation to one another because of their mutual likeness.³⁹

And further,

the beings in question possess certain salient characteristics, are capable of recognizing these in other, similar beings, and acknowledge possession by other beings of the characteristics in question as grounds for following certain rules of conduct toward them.⁴⁰

The implication of such an approach is that non-persons cannot be members of the moral community and therefore one cannot be under any moral obligations towards them — the rules of conduct do not apply to them. There are two reasons to suppose this is a mistake. Firstly, it is not obvious that rules of moral conduct cannot be extended

36 M.A. Fox, "The Moral Community", in H. LaFollette, ed., *Ethics in Practice: An Anthology* (Oxford: Blackwell, 1997), 127–38, at 128.

37 *Supra* n.36, at 129.

38 M.A. Warren, "On the Moral and Legal Status of Abortion", in LaFollette, *supra* n.36, 79–90, at 84.

39 *Supra* n.36, at 128.

40 *Supra* n.36, at 128.

to beings outside the moral community — it may well be correct that only members of the moral community hold rights, but it would be implausible to suggest that only members of the moral community hold interests; and interests can give rise to rules of moral conduct. Secondly, it is not implausible to suggest that non-persons could feature as members of the moral community at some level, and therefore be rights-holders, although obviously the rights they could hold would be limited. This is the implication of Tom Regan's distinction between moral agents and moral patients.⁴¹ Moral agents, he argues, are

individuals who have a variety of sophisticated abilities, including in particular the ability to bring impartial moral principles to bear on the determination of what, all considered, morally ought to be done and, having made this determination, to freely choose or fail to choose to act as morality, as they conceive it, requires.⁴²

Therefore moral agents can be held morally accountable for what they do.⁴³ Moral patients, however, are different:

[A] moral patient lacks the ability to formulate, let alone to bear, moral principles in deliberating about which one among a number of possible acts it would be right or proper to perform. Moral patients, in a word, cannot do what is right, nor can they do what is wrong.⁴⁴

But "moral patients can be on the receiving end of the right or wrong acts of moral agents, and so in this respect resemble moral agents".⁴⁵ Much depends upon whether we see the moral duties towards moral patients as direct or indirect, the latter being duties we have towards them by virtue of duties we directly owe to moral agents. If the duties are direct, as Regan argues they must be,⁴⁶ then it seems arbitrary to claim that moral patients lie outside the boundaries of the moral community. However, even if this is not convincing, and we insist that non-persons remain outside the moral community and therefore rightless, we have already seen that they do hold interests, and that respecting interests can give rise to moral rules of conduct. It follows, therefore,

41 T. Regan, *The Case for Animal Rights* (London: Routledge, 1988), at 151–56.

42 *Supra* n.41, at 151.

43 *Supra* n.41, at 152.

44 *Supra* n.41, at 152.

45 *Supra* n.41, at 154.

46 *Supra* n.41, at 185–94.

that non-persons — moral patients, for instance — can be the subject of such moral rules.

In conclusion, while the concept of a person is indeed a moral concept, it does not follow that the persons/non-persons distinction is a moral distinction; or if it is, then it remains one which is very obscure. On the basis of that moral concept we have to decide how we must act towards persons; but nothing follows about how we must act towards non-persons. In the same way that we were obliged to examine the nature of persons to see what moral conduct was appropriate in their case, we have to examine the nature of particular kinds of non-persons to see what moral conduct is appropriate in their particular case — the "fact" that they are not persons tells us very little. It certainly cannot tell us that to kill a particular non-person does it no moral harm. We must take care that in denying such beings the status of personhood, we do not at the same time deny them the moral value they may nevertheless carry.