Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press, 2002. xiii + 540 pp.

**1.**

This is a book that aims to answer practical questions (such as whether and when abortion and euthanasia are permissible and how we should treat animals and the retarded) by answering such theoretical questions as what we are, when we begin and cease to exist, when it is worth caring about the continuation of our lives, and who is entitled to respect. McMahan provides detailed, rigorously argued, comprehensive, and often unconventional answers to both the theoretical and practical questions. The book is an enormous achievement and required reading for anyone concerned with questions of personal identity, issues of life and death, and the morality governing relations with animals.

The detailed nature of the analysis makes for slow reading in many sections, but never because the text is unclear. McMahan's method of argument relies heavily on intuitive judgments in hypothetical cases. However, he believes that not all our intuitive judgments will cohere and some will simply have to be ignored in formulating a correct theory. (Below I raise some questions about why he chooses to ignore some and not others.) A striking example of this is his willingness to reject the intuitive judgment that a human infant who from conception has only the potential to be severely retarded is no more unfortunate than a normal animal that has the same potential from conception and that is not, he thinks, unfortunate in virtue of this. (Although he thinks—counterintuitively, I believe—that the absence of a better potential is not a misfortune, he holds that the loss of a better potential once had is a misfortune.) Those who would be less willing than he to reject intuitive judgments might argue that the need to ignore some judgments is an indication that we simply have not yet found the correct theory that would accommodate all the judgments.

**2.**

According to McMahan we are embodied minds, and we begin to exist when fetal development reaches the point where the nervous system has the capacity to support consciousness. We cease to exist when the area of our brain that has supported our capacity for consciousness no longer exists or functions. To say that we are essentially embodied minds is not to say that we are essentially persons. By "person" he understands a self-conscious being with some degree of rationality and, apparently, psychological interconnections between temporal stages. We might survive the person-stage of our lives if our mind continues in a demented form.

McMahan also rejects the view that we are essentially organisms; he thinks that our organism began when cells specialized and functioned in an integrated way, but this is not sufficient for our presence, on his account of what we are. He also rejects the view that the early embryo becomes us, since, according to him, while the changes undergone in the transition from an embryo to a late fetus preserve identity of the organism, the organism is not identical to the entity that has capacity for consciousness.

Here are some possible concerns about McMahan's account of what we are. The early embryo is the beginning of our organism, and part of our organism is a brain. (McMahan argues that twins who share the same body from the neck down share the same organism, and yet they are different persons. He concludes from this that persons are not organisms. But this argument seems to ignore the fact that the twins do not have minds that stem from the same brain, and so they do not completely share the same organism.) If the part of the organism that is the brain is the source of the mind, and the embryo is the beginning of an organism that will have a brain, it is not clear why the embryo is not the beginning of us—us under construction—even though it is not yet us (i.e., an embodied mind).

A second type of concern is raised by McMahan's insistence that in order for the same mind to be present, the material substrate of consciousness must remain the same. It is for this reason that he rejects the view that we could survive teletransportation; for the psychology at the other end of the teletransporter would be embodied in a completely different physical material than the original psychology. He recognizes, of course, that normally cells die and are replaced in our brains, but he claims that so long as this happens slowly—in the sense that at any given time new cells are a small fraction of the total cells in the part of the brain that supports consciousness—this is consistent with the same part of the brain giving rise to the same mind. However, if too large a proportion of cells is replaced at a given time, the original embodied mind would not survive, on his account. (It seems that cells could turn over at a very rapid rate consistent with personal identity, so long as they did not turn over in a great mass.) Furthermore, he claims that if at t1 part A of a brain supports consciousness and at t2 part A dies but part B of the same brain supports consciousness—"shining its light" on all the same memories and thoughts once supported by part A—there would be different minds at t1 and t2, and no personal identity over time.

Are these requirements on identity excessively strict? For example, suppose it turned out to have always been true of our brains that the seat of consciousness moves, as cells in a previous area die en masse, with a seamless flow of consciousness throughout. Would we really think that no one had ever survived as long as we had previously thought? Or suppose (counterfactually) that one way that our brains could prevent dementia would be to grow replacements for 75 percent of one's brain cells that had been destroyed by a virus.

Would a particular person who could survive as a mildly demented person with 25 percent of his "original" brain cells be extinguished if such an internal dementia cure took place? Would a particular person be extinguished if we cured dementia by replacing most of the brain cells supporting consciousness using his own stem cells? If so, this would make current research for such a type of cure self-defeating, at least if personal survival is what one is after.

Suppose personal identity were retained in these hypothetical cases. Then what would distinguish them from teletransportation? It seems that if the death of cells in the brain leads the brain to provide new cells that support consciousness there would be personal survival, but when an intervening agency supplies matter that is unrelated to either the original brain cells or the person's own stem cells (as in teletransportation) there would be no survival.

**3.**

Of course, McMahan thinks that whether we survive and whether it matters if we survive are two different questions. He thinks that the presence of properties that account for survival are not sufficient to account for its mattering much to us that we survive. Rather, in cases where we do not split into different branching lives, concern for survival should be a function of what interests we have at a particular time (what he calls time-relative interests) in surviving. These time-relative interests will be a function of (1) the strength of (what he calls) the prudential unity relations (pur) between ourselves at that time and the times we would live through if we survive and (2) the quality of life we would have if we survive. (The most important part of pur depends on overlapping chains of psychological continuity and connectedness between different times of our lives, though the mere survival of the embodied mind provides some pur.) In the absence of any strong pur there is little difference, from the point of view of the interests of an entity, whether it continues or if a new entity appears in its place.

McMahan describes someone who is an "isolated subject," forever under the impression that he has just come into existence and with no thought of his future. McMahan seems to believe that there is no strong reason to care for the sake of an isolated subject that it continue in existence, not because of the inadequate content of each of its present moments but because there is no psychological connectedness and continuity in the life. This seems to imply that there is little more reason for its sake to rescue an isolated subject from death than to rescue an animal whose natural life span is one minute. But if the isolated subject is a self-conscious being who continually thinks that he just came into existence, he can be a person, even if there is little or no psychological connectedness and continuity in his life. This is a synchronic rather than a diachronic conception of personhood. And what if the content of each moment were extremely good and different from other moments? Does the

fact that this subject is not aware of any accumulation of these good moments in his life make his life not significantly more worth preserving for his sake than that of a short-lived animal? I find this hard to believe.

A further concern stems from McMahan's view that a person who could survive (i.e., be the same embodied mind) through psychological changes resulting in no psychological continuity at t2 with his present (t1) state would have reason at t1 to fear being tortured at t2. If significant concern for what condition one will be in if one survives makes sense even without psychological continuity, why cannot there be significant concern for whether one survives—whether one's embodied mind continues—despite the absence of psychological continuity?

If there is reason from the point of view of time-relative interests to care whether one dies, this implies that death can be good or bad for one. But if death involves nonexistence, its goodness or badness cannot be due to death's intrinsic properties, McMahan thinks. Rather he holds that death's goodness or badness for one is due to nonexistence being comparatively better or worse than what would have occurred in the future life with which one would have had pur. McMahan also agrees with the following views (for which I also have argued[1]): (1) often the fact that one's further life could not have involved relevant goods, and so death could not deprive one of them, is what is really bad even if this makes death itself less bad, (2) we should not hesitate to make death itself worse for people if this happens by making their prospects for further goods (with which death can interfere) better, and (3) the badness of a future loss should be discounted by goods one has already had in the past. (In connection with the latter point, the question should arise for McMahan, given his emphasis on pur in evaluating the loss of future goods, whether only those goods in the past with which one has significant pur at the time one would die should be used to discount future losses caused by death. This seems incorrect to me. For suppose someone had undergone a radical psychological change accompanied by amnesia. If pur with the past were important in discounting, then the fact that he had had in the past a long, wonderful creative life would count for very little against the losses he would incur in dying, making his death quite as tragic as that of someone who had had none of these goods.)

Given the view that death is a bad relative to the pur-goods to be had in further life, it is surprising that McMahan does not deal with whether a simple outweighing of pur-goods by pur-bads in further life would make death not be a comparative evil or whether, as Philippa Foot suggested,[2] having certain basic (perhaps pur) goods in future life would be sufficient to make death be a comparative evil. Nor does McMahan consider the possibility (that I dis-

1. In *Morality, Mortality*, vol. 1 (New York: Oxford University Press, 1993).
2. In her "Euthanasia," *Philosophy & Public Affairs* 6 (1977): 85-112.

cussed in *Morality, Mortality*, vol. 1) that the mere fact that a person will be extinguished is an intrinsically evil aspect of death, so that the fact that death involves nonexistence does not mean that it can be only a comparative evil. On account of this other source of the badness of death, one might decide to die in order to avoid the bad things in further life, but only with regret because it means the end of oneself as a conscious being.

**4.**

Can an account of the badness of death provide us with an account of the wrongness of killing, so that wrongness varies with badness? McMahan argues that it does so only for beings who fall below the threshold of ever having been persons. In the case of persons (or past persons), the morality of harm is superceded by the morality of respect for what a person wills and her inviolability. Hence, McMahan's primary concern is to see how the Equal Wrongness Thesis (i.e., the view that it is equally wrong to kill a person who will lose a great deal in dying and a person who will not lose a great deal in dying) can be defended.[3] Since the properties on which respect is based (e.g., rationality, autonomous will) can also come in degrees in different people, a threshold level of these properties must be what is necessary and sufficient for the equal wrongness of killing any person. The truly problematic case for this view, however, is the one McMahan calls the Deluded Pessimist. This involves someone who competently waives his right to life and asks to be killed even though his death would be bad for him, as he is under the mistaken impression that his future life is worse than nonexistence. McMahan thinks both that we would not be showing disrespect for this person's will if we kill him and yet it would be wrong to kill him. He concludes that concern for a person's interests (including not harming him) must be part of respect for that person. Carlos Soto (in unpublished work), however, has pointed out that if concern for interests is a part of respect, this threatens to undermine the Equal Wrongness Thesis, for will it not be less disrespectful to kill someone who would not lose much in dying than to kill someone who would lose a great deal? Perhaps the answer to this problem is that there is a threshold account of interests in the theory of respect comparable to thresholds of rationality and autonomy. Alternatively, perhaps respect for persons is not merely a matter of respecting their choices nor acting from concern for their interests, after all.

McMahan's discussion of the Equal Wrongness Thesis is an instance in which he refuses to ignore an intuitive judgment (that the Equal Wrongness Thesis is correct), even though he raises objections to it and accepts that

3. The view that people are worthy of respect, McMahan says, is different from the idea of the sanctity of life thesis, which makes no reference to rationality and willing. However, McMahan (242) also says that killing a person "is to show contempt for that which demands reverence," and this may blur the difference.

he cannot yet find a theory that adequately accounts for it. He must find his intuitive judgment favoring the thesis to be stronger than the intuitive judgment that always-retarded humans have a different moral status from animals with the same potential. For he is willing to ignore the latter intuition when he cannot find a theory that adequately accounts for it. But some might argue that, for all McMahan says and in the absence of a double standard, there is as much or as little reason to retain the one view as the other.

**5.**

McMahan's views on abortion follow from his views on identity, death, and the wrongness of killing. The conceptus prior to having the capacity for mind is not a subject, and so loses nothing in dying. The later fetus may be an individual subject who loses its future as a person (a stage in the life of an embodied mind), but the individual has very weak pur with that later stage, and so is not harmed much by death. To the extent that one is skeptical about the significance of strong pur in determining how bad death is for someone at a given time, McMahan's account for why one might think that even late abortion and infanticide are not morally problematic will not be convincing.

McMahan, however, is critical of the view that abortion would often be permissible even if the fetus were a person, as some (such as Judith Thomson and I) have argued. He thinks that it is hard to see how abortion could be permissible if (1) one is responsible for a person's having a need for bodily support, even if the person would not have a fate, if his need were not met, that compared unfavorably with never existing, (2) the person is one's biological offspring, and (3) one would have to kill the person to stop providing it bodily support. McMahan admits that he finds the relevance of (1) and (2) to the impermissibility of abortion puzzling; but he joins people who are intuitively drawn to them and their relevance. This is another instance in which McMahan refuses to ignore intuitive responses, despite his inability to find an adequate justification for them, and it contrasts with the way he deals with the intuitive judgment concerning the moral difference between humans and other animals of identical potential. The question is why the standards for accepting (1) and (2) are lower than the standard for accepting this other intuitive judgment.

In my *Creation and Abortion*,[4] I was drawn to the view (similar to McMahan's) that one owes a person that one is responsible for creating more than just her not being worse off than if she had never existed, and more than her just not having a fate that compares unfavorably with never existing. (I was willing to accept this even if the person one created was not one's biological offspring but someone one manufactured.) However, unlike McMahan, I

4. Frances Kamm, *Creation and Abortion* (New York: Oxford University Press, 1992).

thought there were greater limits on what a creator could be morally required to sacrifice in order to see to it that his creation had certain minimal goods, when he had not yet formed a reciprocal personal relation with it.[5]

Furthermore, the emphasis that McMahan places on killing (rather than letting die) in (3) may be wrong. For he allows that it may be permissible to leave one's biological offspring, which one is responsible for creating, to die rather than carry it in one's body. But (as I argued in *Creation and Abortion*) killing someone when he will thereby lose no more than what he gets from the life support to which he has no right merely to save his life, especially when what he gets from the life support causes him to be a threat to the person providing support, may be no more wrong than letting the person die by not providing the life support.

McMahan's views about aborting a person are further complicated by the fact that he thinks that killing a person who is a nonresponsible threat in order to help the person he threatens (even when the latter person is not responsible for the threat existing and abstracting from whether life support is being provided to the threat) is no more permissible than killing an innocent bystander in order to help someone avoid a threat. I find this an implausible view. McMahan supports his view by presenting the Trapped Miner's Case: Due to a shift in rocks, A was hurled against supports that had prevented the collapse of a mine. The collapse reduces the oxygen available to miners in one part of the mine. May these miners kill A (who has enough oxygen in his part of the mine) if only this will make available to them enough oxygen to survive? If we think they may not kill A (the nonresponsible threat who caused the collapse), why should we think that they could permissibly have killed him when he was in the process of being a threat, in order to stop his impact on the supports? And yet it does seem to me that while A is hurtling toward causing the collapse, it would be permissible to kill him if one knew that this alone would stop his impact. The issue at stake here is, I believe, whether (a) the permissibility of stopping the process in which a person would cause harm by harming him implies the permissibility of seeing to it that there is no harm that a person has caused by harming him. (b) but not (a) would license our imposing losses on nonresponsible threats to undo the damage they have caused if these losses could permissibly have been imposed on them to stop their causing the harm in the first place. The (admittedly puzzling) idea is that a process can be bad only because of the harm it will cause, and yet one can make the harm

5. McMahan says that I claim in *Creation and Abortion* that giving up sexual relations is too much to demand of a woman in order to avoid a pregnancy ending in abortion of a person. In fact, I did not claim that giving up sexual relations was too much to demand. I only said that the claim would have to be true in order for an argument for the permissibility of abortion in cases of voluntary sex to be justified. If giving up sexual relations was good, or an easily accomplished task, there would be little reason not to avoid pregnancy that will end in abortion of a person.

not exist only by stopping the bad process itself (of which a person finds himself a part), not by eliminating in other equivalently harmful ways the harm that makes the process bad.[6]

*Frances Kamm*
Harvard University

6. I first discussed this problematic issue in my "The Insanity Defense, Innocent Threats, and Limited Alternatives," *Criminal Justice Ethics* 6 (1987): 61–76, and again in my *Morality, Mortality*, vol. 2 (New York: Oxford University Press, 1996), among other places. I am grateful to Jeff McMahan and Carlos Soto for comments on a draft of this review.

*Philosophical Review*, Vol. 116, No. 2, 2007
DOI 10.1215/00318108-2006-039