Preferences, Death, and the Ethics of Killing[1]

Jeff McMahan

Introduction

If asked to specify an indisputable or paradigm type of misfortune, most people would probably cite death. If asked to cite a paradigm type of immoral action, most people would probably say murder. It seems obvious to most of us, moreover, that there is a strong relation between these two evils, which is that the wrongness of killing is to be explained at least in part in terms of the badness of death for the victim. Indeed, I believe, and will take as an assumption of this paper, a rather stronger claim - namely, that, at least in most cases in which killing is *pro tanto* wrong, the degree to which it is *pro tanto* wrong is a function of the badness of death for the victim. This is because the wrongness of killing is a function of the degree of harm it causes to the victim and death is normally the principal harm that a person suffers in being killed. Killing may, of course, involve other harms to the victim - for example, the violation of his or her autonomy. And there are other factors besides the harm caused to the victim that may contribute to the wrongness of killing or that may affect the extent to which killing is wrong in these cases - for example, side-effects; and also facts about the agent's mode of instrumentality in the occurrence of the death, such as the fact that the agent brings about the harm rather than merely allowing it to occur, that the killing is intended rather than merely foreseen but unintended (or *vice versa*), that it occurs through the creation of a threat rather than through the redirection of a preexisting threat (or *vice versa*), and so on. But, while factors such as these may contribute to determining the degree to which an act of killing is wrong, nevertheless, unless there are further factors (such as guilt or liability) that reduce or nullify the moral significance of harm to the victim, the wrongness of killing will still vary in proportion to the degree to which killing harms its victim.

---

1 This paper was written in great haste in early June 1992 at the request of the conference organizers, who at the last minute found themselves in need of another main paper. As a result, it is often rather rough and sketchy, particularly towards the end. Some parts of this paper are drawn from the manuscript of a book in progress on the ethics of killing. Some of the material that is new to the paper will probably be worked into the book manuscript. Thus comments on any part of the paper would be most welcome.

The question I will address in this paper is what role preferences have in explaining the badness of death. Given the assumption that, other things being equal, the wrongness of killing varies with the badness of death for the victim, the answer to this question should also illuminate the role that preferences have in explaining the wrongness of killing.

The desire-based account

It might be argued that the badness of death can be wholly explained in terms of the frustration of the victim's preferences. Bernard Williams, for example, suggests such a view when he writes that, "if I desire something, then, other things being equal, I prefer a state of affairs in which I get it from [*sic*] one in which I do not get it.... But one future, for sure, in which I would not get it would be one in which I was dead. To want something ... is to that extent to have reason for resisting what excludes having that thing: and death certainly does that, for a very large range of things that one wants. If that is right, then for any of those things, wanting something itself gives one a reason for avoiding death."[2]

Williams is careful to note, however, not only that not all desires are frustrated by death but also that the frustration of some desires by death does not contribute to the badness of death. Desires that are not necessarily frustrated by death are what he calls "non-I" desires - another name for what are known as other-regarding desires or external preferences. Whether the satisfaction or frustration of desires of this sort can contribute to or detract from a person's welfare, or can benefit or harm that person, is of course one of the key considerations in determining whether there can be posthumous benefits or harms.

Desires whose frustration by death does not contribute to the badness of death are what Williams calls "conditional" desires. A conditional desire is one the prospect of whose satisfaction does not provide a reason for continuing to live. The desire to scratch an itch is conditional in this sense: one is indifferent between the satisfaction of the desire and the disappearance of the itch and thus of the desire to scratch. A desire that, by contrast, does provide a reason for continuing to live, and thus in a self-

---

2 "The Makropulos Case: Reflections on the Tedium of Immortality," in Bernard Williams, *Problems of the Self* (Cambridge: Cambridge University Press, 1973), p. 85. Williams considers an alternative account of the badness of death advanced by Thomas Nagel that "does not see the misfortune that befalls a man who dies as necessarily grounded in the issue of what desires or sorts of desires he had." (88) Since he concedes that Nagel's account has some merit, he leaves open the possibility that the non-satisfaction of preferences may not provide a complete explanation of the badness of death.

conscious being naturally gives rise to a second-order desire for the continuation and satisfaction of the first-order desire, is referred to by Williams as a "categorical" desire.[3]

With these distinctions in mind, we may reformulate what might be called the *desire-based account* of the badness of death as follows: the badness of death for a person who dies may be wholly explained in terms of the fact that death prevents the satisfaction of that person's categorical desires about how his own life should go.

An important point to notice about this account is that it presupposes that, for the elimination of a categorical desire by death to contribute to the badness of death, it must be the case that the desire would otherwise have been satisfied - or, rather, there must be a reasonable expectation that the desire would otherwise have been satisfied. For the elimination by death of a categorical desire that would have been frustrated, or would otherwise have remained unsatisfied, had the person lived rather than died is not an instance of death preventing the satisfaction of desire. Death prevents a desire from being satisfied only when the desire would have been satisfied were it not for the intervention of death. (This is not to say that a person who dies when his categorical desires would be doomed to frustration anyway does not suffer a misfortune. He does; but, at least where his desires are concerned, his misfortune is not so much that he dies but that the circumstances of his life were such that there was no prospect of his desires being satisfied even if he had lived.)

The Epicurean objection

Although Williams introduces the desire-based account as a response to the arguments of Epicurus and Lucretius that are intended to show that death cannot be bad for one who dies, this account is vulnerable to an interesting variant of the Epicurean challenge. For it can be argued that what death does is not to frustrate categorical desires but rather to eliminate them; and a desire that no longer exists cannot be frustrated or otherwise fail to be satisfied. To paraphrase Epicurus, when death is, desires are not.

---

3 These categories are not jointly exhaustive of the possibilities. Some non-I desires, for example, are neither categorical nor conditional - for example, a desire for justice in El Salvador. Such a desire is not conditional because just institutions in El Salvador are not something one wants simply on the assumption that one will be alive: one wants these institutions to flourish even if one will not be alive to see it. Nor is such a desire categorical if there is nothing that one can oneself do to promote justice in El Salvador.

I believe that, like the more standard versions of the Epicurean argument, this objection can be met. We may distinguish four possible outcomes, or fates, for a desire: satisfaction, frustration, partial satisfaction, and elimination or disappearance. For present purposes, we will ignore the third possibility. While satisfaction of a desire entails nonfrustration, the possibility of elimination means that nonfrustration does not entail satisfaction. In the case of conditional desires, nonfrustration is normally all that matters. Unless the satisfaction of a conditional desire is instrumental to the satisfaction of a categorical desire [in which case the otherwise conditional desire becomes categorical], it does not matter whether the nonfrustration of the desire is insured through its satisfaction or through its elimination. In the case of categorical desires, however, nonfrustration alone is not enough: it is satisfaction that matters.[4] In cases in which the subject of a categorical desire would persist beyond the extinguishing of the desire, the nonfrustration of the desire through its elimination would clearly not be a satisfactory substitute for its satisfaction. (This is true only of categorical desires that the person endorses at a higher level. If a person has a second-order desire not to have a certain categorical desire - e.g., because he feels it an unworthy desire to have - then its elimination may be better for him either than its satisfaction or its frustration.)

If the elimination of a categorical desire is bad in much the way its frustration is bad when the subject of the desire survives its elimination, then its elimination should also be bad even when the subject is eliminated along with the desire. For the nonsatisfaction of desire is bad independently of the consciousness of it; indeed, the recognition that there are things that may be good or bad for a person other than good or bad states of consciousness is among the insights that have motivated utility theorists to abandon hedonism in favor of preference-based accounts of the good.

The Epicurean may, of course, go on to press the challenge in its more traditional form, arguing that it is not so much the nonexistence of the desire but rather the nonexistence of the subject of the desire that ensures that the nonsatisfaction of the desire that is entailed by death cannot be bad for the

---

4 One inadequate response to the Epicurean objection is to insist that, while one may have no objection to the elimination of a conditional desire, categorical desires are desires that one strongly prefers not to be eliminated. But the objection reemerges at a higher level: for death also eliminates the second-order desire that one's first-order categorical desires not be eliminated.

person who dies. This, however, is not the place to respond to the objection as it occurs in this more general form.[5]

The challenge of marginal cases

There is, however, another, more serious objection to the desire-based account of the badness of death. This is that it is unable to account for the fact that death may be bad for a human fetus and seems inadequate as an explanation of the badness of death for human infants and nonhuman animals.

Consider first the human fetus. My own view, based on the theory of personal identity that seems to me most plausible, is that we are not identical with our physical organisms. Hence there is no necessity to assume that we began to exist at the time that our organisms began to exist - approximately at the time of conception. I believe, by contrast, that each of us began to exist when the brain of his or her organism developed the capacity to support consciousness and mental activity. This occurs sometime during the middle of fetal gestation. Thus, while none of us existed when his or her organism was in the early stages of fetal gestation, we all existed during the later stages of gestation. Because, therefore, a fetus in the later stages of gestation would be one and the same individual as the person into whom it might develop, the fetus suffers the loss of that later life if it dies *in utero*. Death, then, is normally bad for a fetus in the late stages of pregnancy - bad because it deprives the fetus of a life that would be worth living. Yet the fetus clearly has no categorical desires that are frustrated or eliminated by death. Thus, if I am right that death can be bad for a late fetus, the explanation of why fetal death is bad must appeal to more than the frustration or elimination of categorical desires.

A similar point can be made by reference to the case of human infants. While infants have some desires that might be considered categorical (e.g., to be held), they do not have many. Their desires, moreover, are for immediate satisfactions (which is one reason why their desires are so few in number); they cannot, for example, desire that something should be the case tomorrow. Hence the death of an infant at most eliminates only a few categorical desires; and, if death occurs while an infant is asleep, it presumably eliminates no categorical desires at all. Thus, if the badness of death were fully explained in terms of the frustration or elimination of

---

5 See my "Death and the Value of Life," *Ethics* 99 (October 1988), pp. 32-40; and Fred Feldman, *Confrontations With the Reaper* (New York: Oxford University Press, 199_), chs. _-_.

categorical desires, the death of an infant could not be significantly bad for the infant, and might not be bad at all. (It is, I take it, an absurd implication of the desire-based account that the death of an infant may be bad for it if it occurs while the infant is awake but not if it occurs while the infant is asleep.) Yet the death of an infant seems clearly bad - much more clearly bad, say, than the death of a fetus.

Next consider nonhuman animals. They too, like infants, have relatively few categorical desires, all of which, like those of an infant, have temporally proximate objects. So, again, according to the desire-based account, the death of a nonhuman animal could be no worse than the frustration of the few categorical desires (if any) that it might have at the time of its death. It seems, however, that the death of a nonhuman animal is worse than that (as I will argue shortly).

More importantly, notice that the desire-based account appears to imply that the death of a human infant is normally no worse than the death of a nonhuman animal. For, because the infant's temporal horizons are limited in the same way the animal's are, a human infant may have no more, or may even have fewer categorical desires than some nonhuman animal - for example, a dog. But clearly the death of a normal human infant is worse for the infant than the death of a dog is for the dog.

Undesired future goods

What makes the death of a fetus, an infant, or a nonhuman animal bad is primarily that death deprives it of a range of future goods that its life would otherwise have contained.[6] Because these entities lack self-consciousness, they are incapable of foreseeing or contemplating most of these goods and hence are incapable of desiring them. But, because the goods would occur within their own lives were it not for the intervention of death, it seems plausible to regard the loss or deprivation of these goods as a misfortune for them - not just an impersonal loss or loss of impersonal value but a loss that is against the interests of the fetus, infant, or animal itself. Thus the fact that these goods are not and indeed cannot (or cannot now) be desired by the individual in whose life they would occur does not show that losing them through death cannot be bad for that individual.

---

6 Note that this does not apply to a fetal organism whose brain is insufficiently developed to be capable of supporting consciousness and mental activity. The death of such an organism involves a loss, but it is an impersonal loss. It prevents one of us from coming into existence; it does not prevent one of us from continuing to exist.

A similar point can be made about the loss *by a person* of undesired future goods. It is true of all of us that, if we live long enough, our lives will come to contain goods for which we now have no desire. In many cases these are goods that we *would* now desire if we were to envisage them - that is, goods that we do not now desire only because we have not foreseen or contemplated their possibility. The loss of goods of this sort through death clearly seems a personal misfortune despite the absence of desire. [Can the desire-based account be plausibly revised and extended to take these goods into account through counting desires that one would have given relevant ideal epistemic conditions?]

There are also future goods that one does not now desire and would not desire even if one were to entertain the possibility of acquiring or experiencing them in the future. For present purposes let us put aside the question whether there can be "objective" goods that would contribute to the value of one's life even if one did not value them when they occurred, as well as the question whether the loss of future goods of this sort can contribute to the badness of death.[7] Instead let us confine the discussion to goods that one does not now desire even though one envisages them but that one would later be glad to have at the time that one would acquire them. There are two possibilities. One is that the reason that one does not now desire the relevant future goods is that one does not regard them as goods, so that one regards the prospect of coming to desire them or coming to be glad to have them as a prospect of corruption. Although this is admittedly controversial, I believe that the loss through death of goods of this sort can be a misfortune for the person who dies. A longer life, even with intertemporal antagonisms among values, can seem - not least to the possessor of the life during its later stages - to be a better life as a whole than a shorter life whose brevity precludes intertemporal conflicts among values. (If there are objective values - something that a proponent of the desire-based account may be reluctant to concede - then it is of course possible that the values that the person has early in life, and which condemn the values he later has, are mistaken.)

There is also a second possibility, which is that the relevant goods are ones that one now regards with indifference, even though one may recognize that before the time that they would appear within one's life one would

---

7 Thus I will not consider the loss through death of a good that a person previously desired but does not desire at the time of death and would not have valued at the time that the good would have been realized in his life had he lived.

gradually change in such a way as to become glad to have them at that later time. To some, of course, this might seem a psychological impossibility, since it might be thought that, if a rational person were to believe that he will later come to desire some good that he has hitherto not desired, he would then form a desire for that good for the sake of his own future welfare. But, while this may happen in many cases, I can see no reason for thinking that it must happen in all cases.[8] Indeed, it seems that many of life's greatest goods may go unappreciated and therefore undesired until they force themselves upon us. Some people, for example, acquiesce in becoming a parent only to satisfy their partner's longing for a child but then find that the birth of the child brings about a restructuring of their values and concerns such that they come to regard parenthood as more important and satisfying than the goods and activities around which their lives had previously been structured. In this rather common type of case, indifference may survive any amount of reflection, succumbing to replacement by desire only with the actual appearance of the good.

The loss of undesired goods of this sort - that is, future goods that are foreseen or envisaged but to which one remains indifferent - seems clearly to contribute to the badness of death. Indeed, for those who die relatively young (e.g., a teenager who has little concern for the things that would matter to him in middle and old age), the badness of death may be *primarily* attributable to the loss of goods of this sort.

We can now summarize the objections to the desire-based account of the badness of death. The first objection - the Epicurean objection - can be put aside. But the theory seems clearly deficient in its inability to count among the factors that make death bad for those to whom it happens the loss of future goods that are not presently desired. In the case of fetuses, infants, and nonhuman animals, the absence of desire results from the absence of self-consciousness and a consequent inability to desire more than the most temporally proximate goods. In the case of persons, the absence of desire may be the result simply of the failure to anticipate or reflect on certain future possibilities, or it may result from the fact that preferences change over time, so that one may later come to desire that to which one is at present indifferent. But, whatever the explanation of the absence of desire, the loss of presently undesired future goods that would be experienced as

---

8 It is even possible to know that one will later come to desire some good and to regard the prospect of coming to have that desire as an improvement of one's character and yet remain at present indifferent to the good.

desirable at the time they would occur is clearly a factor that contributes to the badness of death. It is, however, a factor that the desire-based account seems unable to accommodate.

It might be argued that the desire-based account can accommodate these goods by extending the range of desires that are relevant. According to the extended desire-based account, death is bad not only because it prevents the satisfaction of existing categorical desires but also because it prevents the formation and satisfaction of future categorical desires. But, while extending the theory in this way might meet the objections posed earlier, it introduces further problems. The main problem is that, if preventing the formation and satisfaction of future categorical desires is bad, then the failure of a person to come into existence will normally be worse, at least in this respect, than the death of an existing person. This is clearly counterintuitive. The obvious move, therefore, is to restrict the account so that it claims only that the prevention of the formation and satisfaction of future categorical desires in the later life of an existing individual is bad for that individual. This may provide all we need for an account of the badness of death while simply evading or deferring consideration of the evaluative issues raised by causing new people to exist.

The problem with this revision, however, is that it is difficult to see how it marks an advance over the insistence that the loss of currently undesired future goods contributes to the badness of death. Indeed, to say that death is bad in part because it prevents the individual who dies from forming and satisfying future categorical desires adds nothing to, and is less perspicuous than, the claim that death is bad in part because it prevents an individual from having future goods that the individual does not now desire but would later value when they would occur. What seems most obviously bad, and explanatorily fundamental, is that, if one dies, one's life will not contain certain goods that it would otherwise contain; that one will also be prevented both from desiring these goods and satisfying one's desires seems of secondary significance.

Return to the type of case mentioned earlier involving intertemporal conflicts among preferences - for example, to the case of a teenager who is not simply indifferent to the things that he would predictably care about in later life but is actively hostile to them. Thus he now prefers not to come to desire those things - indeed, he would prefer to die young (though not right now) than to become the sort of person he will in fact become as he grows older. It would seem that an account of the badness of death that focuses exclusively on the effect death has on the victim's desires would have to hold that the fact that death would prevent this person from forming and

satisfying desires that he now prefers never to have would not contribute to making his death bad for him. It would be odd, in other words, for an account that appeals to preferences alone to give as-yet-unformed first-order desires priority over an actual second-order desire concerning those first-order desires themselves. But, as I suggested earlier, it seems plausible to suppose that the loss of goods that one would have valued had they occurred in one's life, even if one is actively averse to them at the time of one's death, can nevertheless add to the badness of one's death.

[A further challenge to the extended desire-based account might come from an objectivist theory of the good. On this view, certain things are good not because they are desired but are desired because they are good. Hence explanatory primacy of future goods over future categorical desires. But objectivism about the good will presumably be rejected by someone tempted to explain these matters in terms of preferences.]

It seems, therefore, that the badness of death cannot be adequately explained in terms of the fact that death prevents the satisfaction of categorical desires, nor even in terms of this fact together with the fact that death prevents the formation and satisfaction of future categorical desires. The best account of the badness of death must make reference to the loss of future goods that are presently undesired.

The future goods account

One might, indeed, be tempted to conclude that the badness of death is fully explicable in terms of the loss of future goods without reference to preferences at all. The badness of death, on this view, would be entirely a function of the quantity and quality of the future goods of which an individual is deprived by death. Call this the *future goods account* of the badness of death.

This view has an obvious intuitive appeal. And it also has considerable explanatory power. For example, it provides a convincing explanation of why it is normally worse to die earlier rather than later - namely, that the earlier one dies, the more future goods one loses. The future goods account also explains why the death of a human fetus or infant is normally considerably worse for the fetus or infant than the death of a nonhuman animal is for the animal. The explanation is that the future life that the fetus or infant loses would have been both longer and substantially richer in the goods of experience and action than the future life that an animal loses when it dies. Because the infant or fetus's losses are greater, the harm it suffers is correspondingly greater.

Despite these attractions, however, the future goods account is decisively challenged by some of the same marginal cases that embarrassed

the original desire-based account. As we noted, the future goods account implies that, in all cases in which continued life would be worth living, it is worse to die earlier rather than later, since the earlier one dies the more future goods one loses. As we also noted, this seems true in most cases. Thus it is worse to die at 30 than at 90 and worse still to die 15 than at 30. Most of us do not, however, find it plausible that the death of a day-old infant is worse for the infant than the death of a 15-year-old is for the 15-year-old, nor that the death of a seven-month-old fetus is worse than the death of a day-old infant. This is shown by the fact that, while approximately thirty percent of all pregnancies terminate in spontaneous abortion, the level of resources that we as a society are willing to expend to discover ways to prevent spontaneous abortions indicates that even most of those who believe that we begin to exist at conception do not tend to regard these deaths as hidden tragedies. And that we generally regard the death of an infant as less bad than the death of a child or an adult is shown by they fact that it has been routine practice intentionally to allow certain infants born with, e.g., spina bifida, or even Downs syndrome, to die whereas it would be unthinkable similarly to withhold life-saving treatment from a child or adult with one of these conditions. Yet the future good account does in fact imply that, if future life would be worth living and other things are equal, the death of an infant is worse than the death of a child and the death of a fetus is worse than the death of an infant (assuming, of course, that the fetus would be identical with the person who might have developed from it). Indeed, on the future goods account, the worst or most tragic deaths are those that occur immediately after a human being begins to exist (which itself occurs, I believe, sometime during the middle of fetal gestation); yet most of us believe that these deaths are the least bad or least tragic among human deaths that preclude a future life that would be worth living.

The weighted future goods account

Can our intuitions about the comparative badness of death at different ages be defended? I believe that they can and that desires have an important role in the theory that best explains and justifies our intuitions.

Let us say that the badness of death for an individual corresponds to the strength of that individual's egoistic interest in continuing to live. It seems true, as the future goods account implies, that the strength of an individual's egoistic interest in continuing to live depends, other things being equal, on the quantity and quality of the goods that his life would contain were he to continue to live (minus those goods that are, as it were, canceled out by corresponding evils). But other things may not be equal. Consider a day-old infant and a 30-year-old both of whom, if they were to live, would have the

same net amount of good in their lives during middle age. As we have noted, most of us believe that the loss of these goods through death is less bad for the infant than it is for the 30-year-old. So there is something that matters other than the net amount of good that each loses. I believe that the loss of the goods of middle age is less bad for the infant because the infant is less closely psychologically related to itself as it would be in middle age than the 30-year-old is. In general, the strength of an individual's present egoistic interest in some future good is a function of the strength of the psychological relations between the individual now and himself (or herself or itself) at the time at which the good would be realized within his (or her or its) life. (The same is true, *mutatis mutandis*, of the strength of an individual's egoistic interest in avoiding some future evil.)

Before seeking to defend this general claim, it may be helpful briefly to elucidate the notion of psychological relatedness. The following relations are instances of *direct psychological connections*: the relation between an experience and the memory of it, the relation between the formation of an intention and the doing of the act that fulfills the intention, and the relation between an earlier and a later manifestation of a desire, belief, or character trait. When there are direct psychological connections between a person *A* at time *t*1 and a person *B* at *t*2, *A* and *B* are *psychologically connected* with one another. Since the number of such connections may be many or few, psychological connectedness is a matter of degree. *A* and *B* are *strongly psychologically connected* when there are, between them, at least half the number of direct psychological connections that there are between any normal person on one day and himself on the next day.[9]

Suppose that between person *A* at *t*1 and person *Z* at *t*2 there are no direct psychological connections, but that *A* and *B* are psychologically connected, as are *B* and *C*, *C* and *D*, *D* and *E*, and so on down to *Z*. Even though *A* and *Z* are not directly psychologically connected, they are nevertheless linked by an overlapping series of relations of psychological connectedness. Between *A* and *Z* there is thus a relation of *psychological continuity*. Psychological continuity, as I understand it, is a matter of degree. Thus, if the relations between *A* and *B*, *B* and *C*, and so on down to *Z* are relations of strong psychological connectedness, then *A* and *Z* are

---

9 Here I follow Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), p. _. My understanding of the notion of psychological continuity, which is developed in the next paragraph in the text, departs from Parfit's understanding in crucial ways.

strongly psychologically continuous. (It follows that, since *A* and *B* are strongly psychologically connected, they are also strongly psychologically continuous with one another.) If, by contrast, the relations linking *A* and *Z* are overlapping chains of weak psychological connectedness, the *A* and *Z* are only *weakly psychologically continuous*.

When I refer to the degree to which a person now would be psychologically related to himself in the future, what I have in mind is a function of the degree of psychological connectedness and the degree of psychological continuity between that person now and that person as he would be in the future. The more direct psychological relations there will be between a person now and himself in the future, the more closely psychologically related they will be. Similarly, the stronger the relations of psychological connectedness that hold within a life over short periods, the more closely psychologically related the individual at one time will be to himself at another time. In short, the strength of the relations of psychological connectedness and continuity within a life determines the extent to which the life as a whole is psychologically unified.

There are, for example, no direct psychological connections between me now and myself when I was a fetus. And I am only weakly psychologically continuous with myself when I was a fetus, since, during the first year or so of my life (assuming that my life began when I began to exist rather than when I was born), I was only weakly psychologically connected with myself from day to day. Hence the extent to which I as a fetus was going to be psychologically related to myself as I am today is minimal. And this is of course true quite generally: fetuses are only minimally psychologically related to themselves as they will be when they become adults. My claim, then, is that it is because a fetus would thus be only minimally psychologically related to itself later in life that we must discount its egoistic interest in the goods that its life might later contain. Indeed, the fetus's egoistic interest in future goods, its egoistic interest in continuing to live, and the extent to which it would be harmed by losing that life and its constituent goods must all be discounted for the diminished degree to which the fetus would be psychologically related to itself in the future. A similar point holds for infants, though the discount rate is in that case rather less steep.

What reason is there to believe that the degree of psychological relatedness between a person now and himself in the future affects the strength of his present egoistic interest in goods that his life might contain in the future? At the risk of circularity, I suggest that one important reason is that this view provides the best explanation and defense of our intuitive

views about the comparative badness of fetal and infant death. If we discount the badness of the loss of future goods through death for diminished degree to which an individual at the time of death would have been psychologically related to himself at the time the goods would have occurred within his life, then we have a ready explanation of the following intuitions: that the death of a human child or adult is normally worse than the death of an infant, that the death of an infant is normally worse than the death of a fetus late in pregnancy, and that the death of a human fetus is normally worse than the death of a nonhuman animal.

Unless we discount an individual's egoistic interest in future goods for diminished psychological relatedness, we must either give up these intuitions about the comparative badness of different deaths, find another defense of our belief that the death of a fetus or infant is less bad than the death of a child or adult even though it involves a greater loss of future goods, or else devise an alternative account of the badness of death that largely divorces the badness of death from the loss of future goods. None of these alternatives seems more promising than discounting for diminished psychological relatedness.

There is, however, more to be said about the importance of psychological unity. It is the psychological unity within a life that gives the life as a whole a moral and prudential significance that the mere sum of its component experiences lacks - or, to put it differently, that makes the life as a whole a relevant *unit* for moral and prudential evaluation. To see this, consider the hypothetical case of a sentient creature in whose life there are no psychological connections at all. It lives entirely in what is known as the "specious present."[10] That is, while it has enough short-term memory to enable it to see the motion of the second-hand on a clock as a continuous motion, that is *all* the memory it has. It cannot remember what happened to it a few seconds ago. It is not self-conscious and has no conception of the future. Because of the absence of memory, it is incapable of intentional action. Because it is incapable of intentional action, it has no character traits, no dispositions. All it does is have experiences. But most of these experiences are extremely pleasant. Perhaps we can imagine this creature as having a well-formed pleasure center in the brain but little else in the way of neurological development.

---

10 Reference to the case of "Jimmie" in Oliver Sacks, *The Man Who Mistook His Wife For a Hat*.

Would such a creature have an egoistic interest in continuing to live? Could we be concerned *for its sake* that it should continue to live? The mere continuity of consciousness within its life that is a corollary of the functional continuity of its brain seems to provide a basis for the idea that it has some egoistic interest in continued life. But our intuitive sense is that the *egoistic* interest here is minimal. And we can see intuitively that it is the lack of psychological relations between the hypothetical creature now and itself in the future that precludes its having a strong egoistic interest in continuing to live despite the fact that its future experiences would continue to be pleasurable. Each experience in the creature's life is discrete, isolated, unconnected with any of the experiences that precede or follow it. Later psychological states are neither caused by nor make any internal reference to earlier ones. Thus, while we may think that the experiences have value individually, it is less plausible to attribute independent value to them as a collection or aggregate. Lacking any *unity* apart from their common grounding in the same brain, they fail to form a *unit* in any but the most minimal sense. Indeed, it is difficult to think of this series of disconnected experiences as constituting *a life* at all.

Because the experiences fail to form a morally significant unit, it may seem that what matters morally is mainly that experiences of this sort should continue to occur (and even this may not matter much, given the disconnected character of the experiences). It does not much matter whether the experiences occur within this same life. Thus it would not much matter if this creature were to die while another creature of the same sort were to begin to exist and have experiences of the same sort. That which is of most moral significance - the individual experiences themselves - would continue to occur. Nothing of significance would be lost simply because the experiences would be occurring within a different life. For the experiences themselves and not the lives as wholes, or units, would be the focus of our moral concern. The creature itself would be, in Peter Singer's phrase, "replaceable."[11]

---

11 Peter Singer, *Practical Ethics* (Cambridge: Cambridge University Press, 1979), p. 102. The reasoning in this paragraph is influenced by Parfit, who argues that, to the extent that the unity within each life is less deep, the difference between lives will be correspondingly less deep. In other words, "if the fact of personal identity is less deep, so is the fact of non-identity." (*Reasons and Persons*, p. 339) And, when the boundaries between lives are less deep or significant, "it becomes more plausible, when thinking morally, to focus less upon the person, the subject of experiences, and instead to focus more upon the experiences themselves." (341)

If we revise the future goods account to take account of the importance of psychological unity within a life, the resulting theory will have two elements. The badness of any particular death will be determined by [i] the quantity and quality of the future goods (minus future evils) that the individual's life would otherwise have contained, though [ii] the badness of the loss of each future good through death must be weighted for the degree to which the individual at the time of death would have been psychologically related to himself at the time the good would have occurred in his life had he lived rather than died. Call this the *weighted future goods account* of the badness of death.

The relevance of desire

Earlier I claimed that desire has an important role in explaining the badness of death; yet the weighted future goods account makes no reference to desire at all. I will now suggest that, while the weighted future goods account does in fact take some account of the importance of desire, it needs to be further revised in order to take full account of the role of desire in explaining the badness of death.

According to the weighted future goods account, desires have a role in explaining the badness of death because the formation and persistence of a desire constitutes an important form of psychological connection, as does the formation or possession of a desire and its later satisfaction. A life that contains an abundance of categorical desires - especially desires that underlie long-range ambitions, plans, goals, and projects - that persist until they are satisfied is therefore necessarily a life that is richly unified and integrated psychologically. Thus death is worse, other things being equal, the more categorical desires it prevents from being satisfied, since the more the life would have been bound together over time by threads of desire, the worse the loss of each future good becomes.

The weighted future goods account therefore assigns desires a definite though indirect role in explaining the badness of death. A case can be made, however, for the claim that desires also have a more direct explanatory role. Imagine two lives, both highly unified psychologically, that are ended prematurely. In both cases, the futures that are lost would have been roughly equivalent in terms of the net amount of good they would have contained. Suppose, however, that in one case the goods that death prevents would have been primarily ones that the victim strongly desired to have while in the other case many of the goods would have been ones that the victim did not care about while he was alive but would nevertheless have come to value had he lived to experience them. It seems to most of us that, even if both of these individuals would have been psychologically connected to themselves

in the future to a roughly equal extent, the one who is deprived of a future in which the desires he had prior to death would have been satisfied suffers the greater loss.[12]

It seems, then, that desire constitutes an independent third element in the explanation of the badness of death. The loss of some future good through death is worse, other things being equal, to the extent that it was desired at the time of death by the person in whose life it would have occurred. We should therefore revise the weighted future goods account so that it weights the loss of a future good through death not only for degree to which the person who dies would have been psychologically related to himself at the time he would have acquired the good but also for whether, and if so to what extent, he desired the good before he died (and, perhaps, for the length of the time during which he desired it).[13]

There are further complexities here having to do with intertemporal conflicts among desires that I will not take up in detail. There is, for example, a question about the relative badness of the loss through death of a good that the person desires at the time of death but would no longer have desired at the time that the good would have occurred. Perhaps all that need be said at this point is that the loss of a future good is clearly worse if it was desired at the time of death and would also have been desired at the time it would have occurred. The loss is somewhat less bad if the good was not desired at the time of death but would nevertheless have been desired at the time it would have occurred. Finally, the loss of a future good is less bad still if it was desired at the time of death but would not have been desired at the time it would have occurred.

There is at least one other factor that may contribute to the badness of death that is related to the desires of the person who dies. When a person has a long-standing desire for a certain long-term good, he or she is motivated, where possible, to take action to realize that good. Often this involves a considerable investment of time and effort. When death

---

12 I have stipulated, first, that both lives would have been equally highly unified and, second, that one of the person's preferences would have undergone change (so that he would have come to desire the goods that his life would contain). For these two assumptions to be compatible, we must imagine that the life of the person whose desires remain largely constant is less highly unified in certain other ways - for example, his memory may be less good.

13 Again I leave aside the question whether the loss of some future good that was not desired at the time of death is worse if it was previously desired.

intervenes to prevent the satisfaction of a desire that has governed and structured a person's activities in these ways, it thereby retroactively condemns to futility certain significant dimensions of a person's past. This point has been well expressed by Michael Lockwood, who writes that, "set against an ideal of human life as a meaningful whole, we can see that premature death can, as it were, make nonsense of much of what has gone before. Earlier actions, preparations, planning, whose entire purpose and rationale lay in their being directed towards some future goal, become, in the face of an untimely death, retrospectively pointless - bridges, so to speak, that terminate in mid-aid, roads that lead nowhere."[14] (This element of the badness of death is bad in the way that a posthumous harm is: it affects the quality or value of a person's past life for the worse.)

The weighted future goods account should be further revised to take account of this fourth factor. Thus the theory should now hold that the badness of a death for the individual who dies is inversely proportional to the net amount of good that the individual's future life would have contained had he not died, taking into account that the badness of the loss through death of each future good must be weighted for [i] the degree to which the individual at the time of death would have been psychologically related to himself at the time the good would have occurred in his life, [ii] whether, and if so to what extent (and perhaps for how long), the individual desired the good before he died, and [iii] the extent to which the individual's previous activities were dependent for their meaning or value on their being instrumental to the realization of the good.

The desire for continued life

So far the discussion of the role of desires has centered on "local" categorical desires - desires for particular goods within a life. Some writers have assumed, on the contrary, that the badness of death for an individual who dies is to be explained instead in terms of its preventing the satisfaction of the individual's "global" desire for continued life. Peter Singer, for example, argues that a being that lacks self-consciousness and consequently "lack[s] the capacity to desire to go on living" cannot have "a personal [i.e., egoistic] interest in continuing to live," thereby implying that it is the desire

---

14 Michael Lockwood, "Singer on Killing and the Preference for Life," *Inquiry* 22, p. 167. Also see Dorothy Grover, _.

for continued life that grounds the interest in continuing to live.[15] A being that lacks the desire for continued life cannot, according to Singer, be harmed by death or, therefore, by being killed, except insofar as killing it causes it causes it to suffer. He does not, however, conclude from this that killing a being that lacks the desire for continued life cannot be wrong. It may be wrong for impersonal reasons - for example, because it reduces the net amount of good in the world. Thus he writes that, in killing such a being, "one does it no personal wrong, although one does reduce the quantity of happiness in the universe."[16]

Let us call the view that the badness of death can be explained solely in terms of the nonsatisfaction of the desire for continued life the *global desire account* of the badness of death. There are, I believe, at least two reasons why this account is inadequate. First, it is simply implausible to suppose that a being, such as a dog, that does not have the degree of self-consciousness necessary in order to have a desire for continued life cannot have a personal interest in continuing to live. As Singer himself notes, "[a]s long as a sentient being is conscious, it has an interest in experiencing as much pleasure ... as possible."[17] If that is so, and I believe it is, then that being must also have an interest in continuing to live, since it cannot satisfy its interest in experiencing more pleasure unless it continues to live. To put the point more generally, the fact that, like the desire-based account, the global desire account ignores the loss of future goods unless they are actually desired renders it inconsistent with deep and widely shared beliefs

---

15 Peter Singer, Killing Humans and Killing Animals," *Inquiry* 22, pp. 152-3. Compare the discussion in his *Practical Ethics*, p. 100-105. A similar view is advanced by Michael Tooley, who argues that "having a right to life presupposes that one is capable of desiring to continue existing as a subject of experiences and other mental states. This in turn presupposes both that one has the concept of such a continuing entity and that one believes that one is oneself such an entity. So an entity that lacks such a consciousness of itself as a continuing subject of mental states does not have a right to life." See Michael Tooley, "Abortion and Infanticide," *Philosophy and Public Affairs* 2 (1972), p. 49. Also see his *Abortion and Infanticide* (Oxford: Oxford University Press, 1983).

16 *Practical Ethics*, p. 102.

17 *Ibid.*

about the badness of death. Singer himself notes, though understates, its implication that the death of a human infant cannot be bad for it.[18]

A related objection may be more persuasive for those who are tempted to account for the badness of death entirely in terms of the nonsatisfaction of desire. This objection hinges on the fact that the global desire for continued life may or may not be supported or sustained by local categorical desires. Consider, for example, two persons, both with equally intense and long-standing desires for continued life (assuming for the sake of argument that intensity of preference is susceptible of rough measurement). The global desire account implies that death would be equally bad for both. (This assumes that the global desire account does not regard all deaths that are opposed by a desire for continued life as equally bad, but instead allows that the degree to which death is bad for the individual who dies corresponds to the intensity of that individual's desire for continued life.) But suppose that, in the case of one of these two persons, the desire for continued life is based on his numerous local categorical desires - that is, he wants to continue to live *because* he wants to see his children grow to maturity, to finish writing his book, and so on. But suppose that the other person, by contrast, is extremely old and infirm so that the future holds no promise of significant goods for him. Indeed, his life is so full of ills of one sort and another that it is not even clear that any future life he might have would have a positive balance of good over evil. Consequently he has few categorical desires and most of those he has he knows to be doomed to inevitable frustration. Yet, because he is utterly terrified of the prospect of nonexistence, he passionately clings to life and is possessed by a fervent desire to continue to live.

If the badness of death were proportional to the intensity of an individual's preference for continued life, death would be equally bad for these two persons. This, however, seems clearly wrong. Death, for the first of these two persons would prevent the satisfaction of a vast number of categorical desires, whereas it would prevent the satisfaction of only a few categorical desires in the case of the second person. The first person, therefore, has far more to live for; thus his death would be worse.

While the global desire account therefore seems unacceptable as an alternative to the weighted future goods account, it might be possible to devise an *extended desire-based account* that holds that the badness of death for the person to whom it happens is a function both of the number and

---

18 "Killing Humans and Killing Animals,", pp. 153-4.

intensity of the person's local categorical desires whose satisfaction it prevents and also of the intensity of the person's global desire for continued life. Such an account would have an advantage over the narrower desire-based account in that, while the global desire for continued life may in part just be a desire to be able to satisfy one's local categorical desires, it may also go beyond these desires. For one may desire to continue to live in part in anticipation of forming and satisfying future desires whose objects are as yet unpredictable. Thus an extended desire-based account would be able to take some account of the loss through death of future goods that are presently undesired: the loss of these goods is to some extent taken into account by the nonsatisfaction of the global desire for continued life, which may itself be in part a desire to form and satisfy new desires for future goods.

While an extended desire-based account might be thus less vulnerable to the third of the three earlier objections to the original desire-based account, it is, I believe, decisively falsified by the second of those objections. Since neither fetuses, infants, nor most nonhuman animals have a global desire for continued life, the extended desire-based account is no better able than the original, narrower theory to account for the full extent to which death may be bad for fetuses, infants, and animals.

In cases in which death prevents the satisfaction of a global desire for continued life, does this fact contribute anything to the badness of the death that is not already covered by the weighted future goods account? The presence of the global desire does of course establish a presumption that the individual who dies would have been strongly psychologically related to himself in the future had he lived. For in order to desire continued life, a being must be self-conscious to a high degree; and highly self-conscious beings (and *only* highly self-conscious beings) typically have numerous strong, long-term categorical desires as well as a highly developed capacity for memory, belief, intention, and so on - in short, the elements of the psychological connections that are constitutive of psychological relatedness. Yet, while the desire for continued life typically provides evidence of a high degree of psychological relatedness, the importance of psychological relatedness is already fully taken into account by the weighted future goods account.

As previously noted, focusing on the global desire for continued life also provides a way of taking some account within a preference-based framework of the badness of the loss through death of future goods that are not themselves desired at the time of death. But the weighted future goods account does this as well (and does it more directly and perspicuously).

One might, therefore, be tempted to conclude that, in cases in which death prevents the satisfaction of the global desire for continued life, that fact contributes nothing to the badness of the death for the person who dies. This, however, would be a mistake. Return to the case of the person who has virtually nothing to live for but who nevertheless stubbornly clings to life and passionately desires to go on living. Death would, it seems, be bad for this person independently of its depriving him of future goods, independently of its preventing the satisfaction of any local categorical desires, and independently of its preventing the formation and satisfaction of future categorical desires. It would be bad precisely because he wants to go on living. I do not believe that this alone is sufficient to make his death very bad; but it does seem to be a further factor that is not reducible to the other factors we have identified. It should, therefore, be included as a fifth element in the weighted future goods account.

Global desires

I have argued that, insofar as the badness of death is to be explained in terms of the prevention of the satisfaction of desires, it is primarily, though not necessarily exclusively, the nonsatisfaction of local categorical desires, rather than the nonsatisfaction of the global desire for continued life, that contributes to making death bad for one who dies. In an appendix to his book, Parfit has arguments that might be thought to challenge this conclusion.[19] There he argues that versions of what he calls *Desire-Fulfillment Theories* of self-interest that focus on the satisfaction of global desires are more plausible than corresponding versions (which he calls *Summative Theories*) that focus on summing satisfactions (minus frustrations) of both global *and* local desires. Some of his arguments against Summative Theories seem to me mistaken. He argues, for example, that Summative Theories imply that one's life would be better if one were to become addicted to a drug that would have no effects on one's life other than to make one want more of it and that would always be in abundant supply. for the addiction would give one more desires that would then be satisfied, leading to a greater sum of desire-satisfaction over the whole of one's life. The satisfaction each day of a local desire for that day's injection of the drug

---

19 *Reasons and Persons*, Appendix I, esp. pp. 496-9.

would, over time, outweigh the frustration of one's global desire not to become an addict.[20]

Parfit considers, as a response to this objection, the possibility of revising Summative Theories so that they do not take into account either the satisfaction or the frustration of desires that one would prefer not to have. They would then not assign a positive number to the addict's desire for the drug. He shows, however, that this revision is unacceptable, since it implies that Summative Theories cannot count negatively the frustration of a desire not to be in great pain.[21]

There is, however, a more plausible way of revising Summative Theories. The reason that Summative Theories, as Parfit understands them, are vulnerable to the objection based on the case of addiction is that they count the satisfaction of a *conditional* desire positively - that is, they assign its satisfaction a positive number that corresponds to the strength or intensity of the desire. A Summative Theory that distinguishes between conditional and categorical desires and refuses to assign positive value to the satisfaction of conditional desires would not have absurd implications in cases such as that involving drug addiction. Nor need this revision be vulnerable to the sort of objection that Parfit urges against the proposed revision considered in the previous paragraph. For it is perfectly plausible, given the nature of a conditional desire such as the desire to scratch an itch, to suppose that, while the satisfaction of the desire does not count positively (any more than its elimination would), its frustration *does* count negatively. If this is correct, then, while the satisfaction of the desire for an injection would not count positively (i.e., would not improve one's life or make it better), the frustration of that desire would count negatively (i.e., would make one's life worse).

It would, of course, be absurd to treat all preferences in this way. As Michael Lockwood points out, a preference utilitarian analogue of negative

---

20 This argument might be challenged on the ground that the global desire is not just a desire not to *become* an addict but would also be a continuing desire not to *be* an addict. Hence the frustration of the global desire would not be a discrete event but would be continuous, a process. Thus the continuing frustration of the desire not to be an addict might, even over time, outweigh the repeated satisfaction of the desire for an injection. (I know of nothing in the literature on the moral or prudential significance of the continuing frustration of a desire that persists despite its continuous frustration. This would be worth exploring.)

21 *Ibid.*, pp. 497-8.

utilitarianism that counts frustration negatively but does not count satisfaction positively implies that, "unless we can point to some weightier preference, or set of preferences, which would otherwise be frustrated, we should never act in such a way that some preference comes into being only to remain unsatisfied. ... [Hence] it must invariably be wrong to bring new human beings into existence, or to refrain from aborting foetal ones."[22] But the revision of the Summative Theories that I recommend would not have this implication, since it is only the satisfaction of conditional desires that theories would not count positively; the satisfaction of categorical desires would count positively and would thus be capable of outweighing the inevitable frustration of some conditional and categorical desires.

Despite the fact that Parfit's initial arguments fail, he does give one example that successfully supports his contention that a global preference for one future over another may more accurately indicate which of the two futures would be better than a simple determination of which future would contain the greatest sum of weighted local desire satisfaction. His example is the analogue, within a single life, of his well known Repugnant Conclusion. Imagine two possible futures: one that is of normal length but throughout which there is continuous satisfaction of a wealth of intense categorical desires, and another in which there is, each day, a only a marginally positive balance of categorical desire-satisfaction. If the second of these two futures were long enough, the sum of desire-satisfaction it would provide would exceed that provided by the first. Summative Theories would have to conclude, implausibly, that the second life would be better. If, however, we appeal directly to the fact that one would globally prefer the first future, then we get the right answer unproblematically.

This shows that in many, if not all, cases, the relative strength of a person's global preference for a particular future provides a better measure of the value of that future for the person than does the sum of weighted desire-satisfaction it would contain. This conclusion does not, as far as I can tell, affect my contention that the nonsatisfaction of the global desire for continued life contributes only minimally to the overall badness of death. It does, however, suggest the following further conclusion. If the badness of death for the person who dies is proportional to the value, for that person, of the future life that death prevents him from having, then we may get a more accurate assessment of the badness of his death by seeing what global desires (e.g., about what sort of life the person wants or what sort of person

---

22 Lockwood, *op. cit.*, p. 164.

he wants to be) death prevents from being satisfied than by attempting to determine the sum of weighted satisfactions of the person's local categorical desires that death prevents.

This concession may require father modifications of the weighted future goods account.[23] [The value of a future may be a function of the complex interplay of various factors, including global preferences (which may themselves be distorted or irrational and hence not always a reliable gauge of the value of a future), the sum of local categorical desire-satisfaction, the net goods the future would contain, appropriately weighted for psychological relatedness, etc.] I will not, however, attempt to make further modifications here. I will instead conclude by sketching some of the implications of the weighted future goods account for the ethics of killing.

Abortion, infanticide, and the killing of animals

Assuming that the degree to which killing is *pro tanto* wrong is a function of the amount of harm that it causes to the victim, we can appeal to the weighted future goods account of the badness of death to explain why the killing of fetuses, infants, and nonhuman animals is normally less morally objectionable than the killing of normal adult human beings. Let us begin with the killing of fetuses. Assume for the sake of argument that abortions involve killing the fetus.[24] If we do not begin to exist until the brains of our fetal organisms develop the capacity to support consciousness and mental activity, then an abortion performed prior to that point does not kill one of us (though it kills an organism) but instead prevents one of us from coming into existence. Apart from considerations of side-effects, such an abortion is morally on a par with an act of contraception. (For present purposes I leave it an open question what the morality of preventing the existence of people who would have lives worth living is.)

Once one of us begins to exist in association with the fetal organism, abortion then harms that individual by depriving it of a future life that would have been its own. But, while the sheer quantity of future goods of which the later fetus is deprived by death is considerably greater than that of which

---

23 Imagine a future life in which there is a tiny net sum of good every day. And imagine that this future would be infinite. Surely its value would not be infinite. Can this be explained solely in terms of the fact that goods in the distant future would have to be discounted to zero for diminished psychological relatedness?

24 In "Killing, Letting Die, and Withdrawing Aid" (forthcoming in *Ethics*) that not all abortions involve killing.

death deprives an adult, this fact is offset and indeed outweighed by the fact that the fetus's egoistic interest in the goods of which it is deprived must be heavily discounted for the diminished degree to which the fetus would have been psychologically related to itself in the future when it would have acquired the goods. In particular, the harm that the late fetus suffers in being killed must be discounted for the fact that its death does not involve the nonsatisfaction of categorical desires, nor the nonsatisfaction of the desire to live, nor the retroactive nullification of previous efforts or sacrifices intended to improve the future that is lost. We may conclude that, while there is a moral objection to late abortion based on the harm that the later fetus suffers in being deprived of future goods, the objection is considerably weaker than that which applies to the killing of a normal adult human being.

Granting this, we must still ask: *how much* weaker is the objection to late abortion? For late abortion might be considerably less objectionable than the killing of a normal adult human being and yet still be seriously wrong. We cannot, of course, expect precision here. But there is a way of consulting our intuitions that may help us to determine roughly how much less objectionable late abortion is than the killing of an adult. Suppose we ignore for the moment the fact that a significant part of the objection to the killing of an adult is that this violates the individual's autonomy - an objection that does not apply in the case of abortion. Then the wrongness of killing the adult will be proportional to the badness of death for that adult. Since the wrongness of killing a late fetus is also proportional to the badness of death for the fetus, the killing of the fetus should be less objectionable than the killing of the adult to roughly the same extent that the badness of death for the fetus is less than the badness of death for the adult. Since most of us regard the death of a late fetus as very considerably less tragic than the death of an adult, we may infer that late abortion must be very considerably less objectionable than the killing of an adult. If we now reintroduce the fact that the killing of an adult is also objectionable on the ground that it violates the victim's autonomy, then the divergence between the wrongness of killing an adult and the wrongness of late abortion becomes even greater.

Given the way in which we in fact develop, the psychological ties between an individual and himself in the future are weakest in the period immediately following the beginning of his existence. Indeed, for a certain period following the beginning of an individual's existence, there may be no psychological ties at all binding him to himself in the future. For this reason, abortion just after the late fetus begins to exist may be only minimally morally objectionable. (Because of this, it may not make much difference morally whether we can locate with any precision the period

during fetal gestation during which we begin to exist - i.e., the time at which the fetal brain acquires the capacity to support consciousness and mental activity.) This squares with our intuitive sense that there is no point in the process of gestation at which abortion suddenly becomes seriously wrong, where only a short time earlier it would have been relatively unobjectionable. As the process of gestation progresses, the late fetus's psychological capacities develop and expand and the unity and complexity of its mental life increase correspondingly. As it thus becomes more closely psychologically linked to its own future, death becomes increasingly harmful to it. Thus the moral presumption against abortion, which originates or begins to apply when the hemispheres develop the capacity to support consciousness and mental activity, becomes increasingly stronger as gestation progresses toward birth. This supports and helps to explain the common view that abortion becomes morally a more serious matter the later in pregnancy it is performed.

This does not mean, however, that the morality of abortion changes drastically between the time that the hemispheres develop the capacity for consciousness and the time of birth. It does not mean, for example, that late abortion changes from being relatively innocuous to being seriously wrong just prior to birth. That would be true if the degree to which the later fetus would be psychologically related to itself in the future were to increase dramatically over the last few months of gestation. But there is no evidence that such a dramatic change in the mental life of the fetus in fact occurs. There is, however, reason to believe that quite a significant change does occur at birth. For at birth the infant's mind begins to be bombarded with stimuli which impel it to a higher level of activity, causing it to form increasingly more psychological connections from day to day. Hence the harm that an infant of only a few weeks may suffer through death may be significantly greater than that which is suffered by a late fetus.

This fact helps to explain and justify the common view that infanticide is morally more objectionable even than very late abortion. (There are, of course, other elements in the complete explanation, such as the fact that, once the child can be seen and felt, people tend to form emotional bonds with it that were not possible when it was confined in the womb. And we regard the death of one to whom we are closely emotionally attached as worse than the death of one to whom we are not similarly attached.) Yet, while infanticide is, on this analysis, more objectionable than abortion, it remains less objectionable than the killing of an adult. This is because, even though an infant is more closely psychologically related to itself in the future than a late fetus is, the psychological connections between an infant and the

person it might become are still extremely weak. Moreover, while it is not wholly implausible to suppose that an infant has some desires that might be considered categorical, the death of an infant does not frustrate more than a few categorical desires. And because human infants are not self-conscious and do not engage in goal-directed activity, the death of an infant cannot frustrate the desire to continue to live or condemn to futility any future-directed activities from the infant's past. Finally, early human infants have no more capacity for autonomous choice and action than the later fetus has, so the harm involved in the violation of autonomy is also absent in the case of infanticide.

Finally, let us turn to the killing of nonhuman animals. The weighted future goods account recognizes that even non-self-conscious animals can be harmed by being killed, since killing them may deprive them of a future that would have contained a net balance of good. Yet the killing of a nonhuman animal is normally less harmful to the animal even than the killing of a human fetus is to the fetus. For, while it is presumably the case that an animals would be more closely psychologically related to itself in the future than a fetus would be, the future life it would lose through death would be incomparably poorer in the goods of experience and action than the entire life of a normal human being.

So the goods that an animal's future life would contain were it not for the intervention of death are severely limited in both quantity and quality. Furthermore, since most nonhuman animals lack self-consciousness, or are self-conscious only to a limited degree, their lives lack a significant degree of psychological unity. Hence the harm they suffer in being deprived of future goods by death has to be discounted for this diminished degree of psychological relatedness. This harm must also be discounted for the fact that the relevant goods are not usually ones that the animal desires prior to its death. And because of this, it is not the case that the value of an animal's previous life can be diminished by death's robbing its past activities of their meaning.

According to the weighted future goods account, together with the assumption that the wrongness of killing is a function of the harm that killing causes to the victim, the killing of an adult person will normally be worse than the killing of an infant, the killing of an infant will be worse than the killing of a late fetus, and the killing of a late fetus will normally be worse than the killing of a nonhuman animal. This, I think, coincides with what most of us believe.