Death and the Value of Life*

Jeff McMahan

Among the services that philosophers have traditionally attempted to provide is the manufacture of arguments intended to show that death is not or cannot be bad for those to whom it happens. In the first section of this paper I will contend that the most influential of these arguments fails to establish the conclusions which its defenders have sought from it. I will then devote the bulk of the paper to developing an account of why it is that death can be bad for those who die. Finally, I will sketch an apparent paradox which threatens this account and conclude by proposing a way of dissolving the paradox.

I. THE EPICUREAN ARGUMENT

A. The Existence Requirement and the Wide and Narrow Experience Requirements

The argument I will consider derives from Epicurus.[1] Death, it is claimed, cannot be bad, or be a misfortune for the person who dies, for, when death occurs, there is no longer a subject to whom any misfortune can then be ascribed. This of course assumes, as I will assume throughout this paper, that death consists in the annihilation, or the ceasing to exist, of the person who dies.[2]

* I have been helped in writing this paper by comments on earlier drafts by Matthew Buncombe, William R. Carter, Gerald Dworkin, Dorothy Grover, Thomas Hurka, Gregory Kavka, Brian Pike, Alan Mattlage, Robert McKim, Richard Mohr, and Jan Narveson. And I am greatly indebted to David Brink, Hugh Chandler, Timothy McCarthy, Derek Parfit, and Steven Wagner for exceptionally helpful comments and discussion. For discussion both of certain aspects of the problem of death which are not addressed in this paper and of certain objections that might be raised to the arguments developed here, see chap. 1 of my book, *The Ethics of Killing* (Oxford: Basil Blackwell, 1989).

1. Epicurus, "Letter to Menoeceus," 124b–127a, in *Letters, Principal Doctrines, and Vatican Sayings,* ed. Russel M. Geer (Indianapolis: Bobbs-Merrill, 1964).

2. Strictly speaking, death is a biological phenomenon—something that happens to living organisms. Since I believe that persons are not identical with their organisms, I believe that persons can cease to exist in other ways than through death (e.g., a person will cease to exist if he loses consciousness irreversibly, and there are cases in which this is compatible with his organism's continuing to exist and even continuing to live); and, though this never in fact happens, it is in principle possible for persons to survive the deaths of their organisms. As things are, death is a sufficient condition of the ceasing to

Ethics 99 (October 1988): 32–61
 0014-1704/89/9901-0007$01.00

Call this the Epicurean argument. It presupposes the following principle:

> *The Existence Requirement.*—A person can be the subject of some misfortune only if he exists at the time the misfortune occurs.[3]

The Epicurean argument should be distinguished from a similar argument which is based on a principle that Epicurus seems also to have held—namely, the principle that an event can be bad for a person only if he experiences it as bad. Call this principle the *Narrow Experience Requirement.* It implies that, while the prospect of death can be bad for a person if the anticipation of it disturbs him, death itself cannot be bad for him since, when it occurs, he will not experience it and, a fortiori, will not experience it as bad.

The Narrow Experience Requirement should in turn be distinguished from the *Wide Experience Requirement,* which holds that an event can be bad for someone only if it in some way affects or makes a difference to his conscious experience. The difference between the two principles is that, while the Narrow Requirement implies that death cannot be bad for the person who dies, the Wide Requirement does not. For, although one does not experience death, it does affect one's experience—by limiting or ending it.

Obviously, then, to sustain the belief that death can be bad for the person who dies one needs to reject only the Narrow and not the Wide Experience Requirement. Yet the two are frequently conflated, so that it is often assumed that one must reject the Wide Requirement as well. This is a natural mistake, since the rejection of the Wide Requirement entails the rejection of the Narrow Requirement. Moreover, although the rejection of the Narrow Requirement does not entail the rejection of the Wide Requirement, the main objections to the Narrow Requirement also apply with equal force to the Wide Requirement. For example, the hypothetical cases in which people who long to be loved and admired and believe that they are, when in fact they are ridiculed and despised behind their backs, constitute counterexamples to both requirements.[4]

In my view these counterexamples provide sufficient grounds for rejecting the Narrow Requirement. But I will not press the objection

---

exist of a person, but not a necessary condition. In this paper my concern is with the ceasing to exist of persons. But for convenience I will treat the death of a person as if it were equivalent to the ceasing to exist of a person.

3. The Existence Requirement does not, of course, require that the person exist when the cause of the misfortune occurs.

4. See, e.g., Thomas Nagel, "Death," in *Mortal Questions* (Cambridge: Cambridge University Press, 1979), p. 4; and James Rachels, *The End of Life: Euthanasia and Morality* (Oxford: Oxford University Press, 1986), pp. 46–48. An equally damaging objection to both Experience Requirements is that they have trouble accounting for the fact that experiential states such as happiness and unhappiness can be rational or irrational, appropriate or inappropriate. See Jeff McMahan, *The Ethics of Killing* (Oxford: Basil Blackwell, 1989), chap. 1.

here, since my target is instead the Epicurean argument. I mention the Narrow Requirement mainly to ensure that it is not confused with the Existence Requirement. To see how the two requirements diverge, consider the case of a person on holiday on a remote island. Back home, on Friday, his life's work collapses. But, because of the inaccessibility of the island, the bad news does not arrive until the following Monday. On the intervening Sunday, however, the man is killed by a shark; so he never learns that his life's work has come to nothing. While it is compatible with the Existence Requirement to suppose that he suffered a terrible misfortune when the collapse occurred, since he existed at the time, the Narrow Experience Requirement implies that this was not a misfortune for him since he was never aware of it. The Existence Requirement and the Narrow Experience Requirement are therefore quite different requirements, and the counterexamples that are often advanced against the latter leave the former, and therefore the Epicurean argument itself, entirely unscathed.

B. The Reconciliation Strategy

The obvious attraction of the Epicurean argument is that it is thought to have the welcome effect of undermining the idea that death is something to be feared. It is less often noticed, however, that, for the same reasons, the argument also threatens certain other commonsense beliefs that we are less eager to abandon—for example, that killing is wrong, that suicide can be rational or irrational, and so on. If we wish to salvage these latter beliefs, we must either reject the Epicurean argument or else show that the relevant beliefs are in fact compatible with the argument's conclusion. If we adopt the first option then obviously we will lose our ground for thinking that death should not be feared. I will argue that the most plausible way of pursuing the second option also has this consequence.

Though the two options are mutually exclusive, I will argue for them both. While I think the Epicurean argument should ultimately be rejected, the other option—which I will call the *reconciliation strategy*—is worth exploring as a fallback position in case the proposed reasons for rejecting the Epicurean argument prove unpersuasive.

Let us consider the reconciliation strategy first. It can be introduced by drawing on a comparison between death, or the ceasing to exist of a person, and the coming into existence of a person. It can be argued that, if a person's life would be worth living, then it would be good—not just impersonally, but good *for that person*—if he were to come into existence. The argument for this claim derives from the compelling commonsense view that it would be wrong, other things being equal, to bring a person into existence if his life could be expected to be utterly miserable. The most plausible explanation of why this would be wrong appeals to two claims: first, that to cause such a person to exist would be bad for that person, or would harm him; and second, that there is a general moral reason not to do what would be bad for people, or would harm them.

But, if to be caused to exist with a miserable life can be bad for a person, then it should also be the case that to be caused to exist with a life that is worth living can be good for a person.

To deny this—to accept that to be caused to exist can be bad for a person but to deny that it can be good for a person—would, in the absence of some explanation, be unacceptably ad hoc. And it is difficult to imagine what sort of explanation could be given for the claim that there is an asymmetry of this sort. Moreover, the moral implications of supposing that there is such an asymmetry are likely to be quite implausible. For example, if it were possible to harm but not to benefit people by bringing them into existence, then any decision to have a child would carry a risk of harming the child but would not involve a possibility of benefiting him. Thus there would always be a moral presumption against having children. But it seems clear that there is no general presumption of this sort. So it seems safe to assume that, if causing a person to exist can be bad for that person, then it can also be good for him. It is, of course, possible to deny both these claims, but then one would need to find an alternative explanation of the fact that it would be wrong to cause a miserable person to exist.[5]

The point that emerges from this that is relevant to the reconciliation strategy is that, if we bring a person into existence with a life that is worth living, there is no problem in locating the beneficiary. He is here before us, enjoying the goods of life. There is thus no obvious incompatibility between the claim that it can be good for a person to come into existence and the Existence Requirement. There is, moreover, a relevant parallel here between starting to exist and continuing to exist. You might have died five years ago. But you did not; instead you continued to exist. If your life has been worth living, then it was good for you to have continued to exist. There is no problem in this case in locating the subject of this good fortune: you are the beneficiary of your continued existence, and you clearly exist during the time when this good fortune occurs. Hence it is not obviously incompatible with the Existence Requirement to claim that your continuing to exist is a good thing for you.

The claim of the reconciliation strategy is that continuing to live can be good for a person even if, as the Existence Requirement implies, death would not be bad for him. There is, however, a way in which the Epicurean might try to establish that the claim that continuing to live

5. One alternative possible explanation appeals to the idea that it is wrong, other things being equal, to do what makes the outcome worse in impersonal terms by increasing the net amount of whatever it is that we regard as bad—suffering, for instance. Thus we might object to causing a miserable person to exist on the ground that this produces an uncompensated increase in the amount of suffering in the world. If we appeal to this alternative explanation, then we will be able to develop a way of reconciling the conclusion of the Epicurean argument with the various commonsense beliefs which it threatens that is analogous to the way developed in the text. The difference is that, according to the alternative account, death will be impersonally bad rather than what I call quasi-impersonally bad. (See n. 8 below.)

can be good is incompatible with the claim that death cannot be bad. This involves appealing to what has been called the "Comparative View." The Comparative View is the view that any judgment about what is good or bad for a person must be implicitly comparative in the following way. The judgment that an act or event is *good* for someone implies that, if there is no relevant alternative that would be equally good or better, then the relevant alternative or alternatives to that act must be *worse* for that person. Similarly, the claim that an act or event is *bad* for someone implies that, unless there is some relevant alternative that would be as bad or worse, then any relevant alternative must be *better* for that person.[6]

The conclusion of the Epicurean argument—that death is not bad for those who die—together with the Comparative View, implies that the alternative to death—namely, continuing to live—cannot be good. This implication, if allowed to stand, will defeat the reconciliation strategy. Thus to defend the reconciliation strategy, we must reject the Comparative View.

One objection to the Comparative View is that it seems to lead to paradoxes.[7] Consider some person who continues to exist. The Existence Requirement itself (which for the moment we are treating as a fixed point in the argument) provides no ground for denying that continuing to exist is good for this person. And if continuing to exist is good for the person, then, given the Comparative View, it follows that ceasing to exist would have been worse for him. But suppose now that the actual outcome is that the person dies. According to the Existence Requirement, this is not bad for the person. Given the Comparative View, it then follows that it would not have been good, or better, for the person had he in fact continued to exist. A similar paradox arises when we consider the alternatives of coming into existence and not coming into existence.

The Comparative View thus has conflicting implications depending, it may seem, on which outcome we suppose to be the actual one. The Epicurean may reply, however, that the paradoxes derive, not from the Comparative View itself, but from our failure to evaluate the outcomes consistently. If (he might argue) we think that ceasing to exist cannot be bad, then we must accept that continuing to exist cannot be good, and we cannot change our minds about this when the actual outcome is continued existence rather than death. If we consistently hold that neither coming to exist nor not coming to exist is either good or bad, and that neither continuing to exist nor ceasing to exist is either good or bad, then the Comparative View will not lead to paradoxes.

The paradoxes to which I have called attention do not, therefore, immediately undermine the Comparative View. They do, however, help

6. The Comparative View is both christened and advocated by Jan Narveson in "Future People and Us," in *Obligations to Future Generations*, ed. Richard Sikora and Brian Barry (Philadelphia: Temple University Press, 1978), p. 48.

7. Here I draw on the argument of my "Problems of Population Theory," *Ethics* 92 (1981): 104–5.

us to understand more clearly what is at issue here. The Epicurean starts with the Existence Requirement, infers that death cannot be bad for those who die, and then reasons via the Comparative View that continuing to exist cannot be good. The proponent of the reconciliation strategy starts with equal propriety with the claim that in most cases continuing to exist is good and then, seeing that an inconsistency arises when this claim is conjoined with both the Existence Requirement and the Comparative View, concludes that the Comparative View must be wrong. Thus to resolve this dispute, we must assess the relative plausibility of the Comparative Requirement and the claim that continuing to exist can be good.

As I suggested earlier, the claim that continuing to exist can be good is supported by the parallel between coming to exist and continuing to exist. I argued that it can be good for a person to come into existence and concluded that, if this is so, it can also be good for a person to continue in existence. This is, of course, a relatively weak case; no doubt more could be said. But for our purposes even a weak case will do. For there is no case for the Comparative View. The Comparative View takes a truth about standard cases of something being good or bad for a person and generalizes it, making it a necessary condition in all cases for something to be good or bad for a person. In standard cases, in which the person exists in both alternative outcomes, the Comparative View seems true as a matter of logic. But in the two nonstandard pairs of alternatives (coming to exist or not coming to exist and continuing to exist or ceasing to exist) in which the person exists in only one of the alternatives, there is no reason whatever to think that the Comparative View must be true. There is no reason to generalize from the standard to the nonstandard cases.

Since the Comparative View is without foundation in the two nonstandard cases, it should be rejected as a general necessary condition for something being good or bad for a person. The rejection of the view then allows us to reject the idea that never existing can be bad or worse for a person, while accepting that to come to exist can be good for a person. We can say that coming to exist can be good, even though it is not better than never existing—since the latter implies that never existing would be worse, which it would not be (at least if it were the actual outcome). This also allows us to concede the Epicurean claim that death is not bad for people, or worse for them than continuing to live, while at the same time accepting that continuing to live can be good for them. We can say that continuing to live can be a good thing for a person, though it is not better than dying. In short, we simply give up the comparative claims, while retaining the noncomparative ones. (If one insists on having a comparison, one could claim that, while never existing would not have been worse for me, it would not have been as good for me as existing has been, and similarly that, while dying would not be bad for me, it would not be as good for me as continuing to exist would be. These claims are true and unparadoxical, and therefore reinstate a com-

parative view of sorts, though not one which threatens the reconciliation strategy.)

Suppose we accept that continuing to live can be good for a person even if death would not be bad for him. We could then claim that death would be bad—not bad for the person who dies, but bad in a quasi-impersonal way.[8] It would be bad because it would exclude what would be good for a person—namely, continuing to exist. Death would be bad in the same way that it is bad if a person whose life would have been worth living fails to come into existence. In both cases something which would be good for a person fails to occur, though in both cases the nonoccurrence of the good is not bad for the person who would have experienced it, since the nonoccurrence of the good involves—indeed perhaps consists in—the nonexistence of the person.

If death can be bad in this quasi-impersonal way, then this provides a sufficient basis for the important beliefs that seem to presuppose that death can be bad for the person who dies. If, for example, death can be bad because it excludes what would have been good for a person, then this will provide a foundation for the belief that killing is wrong, the belief that it is not irrational to fear death, the belief that suicide can be irrational, and so on.[9] In this way these important commonsense beliefs can be reconciled with the Epicurean claim that death cannot be bad for those who die.

C. The Existence Requirement

Despite the apparent success of the reconciliation strategy, it may seem that a more satisfactory way of preserving the relevant range of commonsense beliefs would be simply to reject the Epicurean argument. And this argument can be challenged directly. Consider again the case of the person whose life's work collapses while he is on holiday on a remote island. Suppose we agree that the fact that his life's work has come to nothing is a misfortune for him.[10] On reflection, it seems hard to believe that it makes a difference to the misfortune he suffers whether the collapse of his life's work occurs shortly before he is killed or shortly afterward. Yet, according to the Existence Requirement, this difference in timing makes *all* the difference. If the collapse of his life's work occurs just before he dies, then, even though he never learns of it, he suffers a

8. I call the badness of death as it is here understood "quasi-impersonal" to distinguish it both from the badness of events that are bad because of their effects on people and from the badness of events which are bad in a fully impersonal way—that is, bad for reasons that are independent of considerations of effects on particular people.

9. One could appeal to a parallel reconciliation strategy to show that suicide can also be rational—e.g., in cases in which death would be quasi-impersonally good, excluding what would be bad for a person.

10. The Epicurean, it seems, can reject this claim only by appealing to the Narrow Experience Requirement. He cannot appeal to the Wide Requirement, since this is incompatible with the Existence Requirement.

terrible misfortune. If, on the other hand, it occurs just after he dies, he suffers no misfortune at all. If we find this hard to believe then we may be forced to reject the Existence Requirement.

Most of us, however, will be disposed to do this anyway. The example is probably superfluous, since most of us find death itself a sufficient counterexample. Apart from suffering great pain, it is hard to think of a clearer example than death of something that most people believe to be in most cases bad for the person to whom it happens. The Epicurean simply denies what most of us believe. Death, he claims, cannot be bad for us because when it occurs we will not exist at all. But that is precisely what we object to: that we will not exist when we might otherwise be enjoying the benefits of life.

Death differs from never existing in one crucial respect. Never existing is not something that ever happens to actual people. A fortiori, there are no actual people for whom never existing can be bad. But death always happens to actual people. It can deprive actual people of what would otherwise be good for them. The Existence Requirement stands in the way of our being able to conclude from this that death can be bad for the actual people to whom it happens. But what authority does this principle have to overturn an evaluative inference so fundamental and compelling as this? The conflict here is analogous to a type of conflict that commonly arises in moral theory. For all of us there are certain convictions which constitute more or less fixed points in the system of our moral beliefs. When a conviction of this sort clashes with the dictates of some moral theory, the challenge from the theory must be more rationally compelling than the conviction itself if the conviction is to be justifiably dislodged. Theories seldom satisfy this demand. In the present case, the Existence Requirement also lacks the requisite rational force. Indeed, it seems, like the Comparative View, to be a simple misgeneralization from standard cases to an importantly different nonstandard case. In most instances it is a necessary truth that a person must exist to be the subject of some misfortune: I cannot, for example, suffer the pain of a toothache unless I exist. But death is obviously a special case. To insist that it cannot be an evil because it does not meet a condition that most if not all other evils satisfy is tantamount to ruling it out as an evil simply because it has special features.

It is important to notice, however, that if we reject the Existence Requirement, we may get the claim that death can be bad for those who die as part of a larger package which includes claims about posthumous benefits and posthumous harms. Death, when it is an evil, is a privative evil. In itself it has no positive or negative features, so when it is bad this must be because of what it deprives us of. But many people believe that people who have ceased to exist may also be subject to positive evils and misfortunes—as when a person's life's work is unjustly despised or neglected after her death—as well as positive benefits, as when an artist's works achieve the recognition they deserve which they never received

during her lifetime. It seems that the only way we can reject both the Existence Requirement and the idea that we can be posthumously harmed and benefited is to accept some view, such as hedonism, which incorporates the Wide Experience Requirement. If we are hedonists we can believe that death is bad for the person who dies because it deprives him of valuable mental states, but we can reject the notion of posthumous harms and benefits on the ground that a dead person's mental status can no longer be affected. If, however, we do not accept the Wide Experience Requirement, then the rejection of the Existence Requirement would seem to commit us to the idea that numerous events that occur after our deaths can be either good or bad for us, or in or against our interests.

II. THE BADNESS OF DEATH

Given, then, that death can be bad, either for the victim or in quasi-impersonal terms, how are we to understand the loss or deprivation involved in death? The correct general answer to this question would seem to be that given by Nagel, which is that death, when it is bad, is bad because it deprives us of possible future goods.[11] The apparent simplicity of this answer, however, conceals a host of difficult problems. This section will be devoted to exposing the nature of these problems and proposing solutions to them.

One such problem is implicit in the following remarks of Nagel's. "Countless possibilities for [an individual's] continued existence are imaginable, and we can clearly conceive of what it would be for him to go on existing indefinitely. However inevitable it is that this will not come about, its possibility is still that of the continuation of a good for him, if life is the good we take it to be. . . . Death, no matter how inevitable, is an abrupt cancellation of indefinitely extensive possible goods. . . . If there is no limit to the amount of life that it would be good to have, then it may be that a bad end is in store for us all."[12]

If, however, the goods that would be possible for us were it not for death are conceived of as potentially unlimited, then it may seem that there is nothing to prevent the conclusion that our losses in death are infinite. But then, if the loss involved in death is infinite, and if the badness of death consists in what it deprives us of, how can we explain the common and compelling belief that it is in general worse to die earlier rather than later—for example, that it is worse, or more tragic, when someone dies at thirty than it is when someone dies at eighty?

11. Nagel, pp. 7–10. If, as I suggested in Sec. I.C., there can be posthumous benefits, then death does not deprive us of all possibilities for good. It deprives us only of possibilities for goods of an active or experiential kind. This and other related points are made by Frances Myrna Kamm in her *Morality, Mortality* (Oxford: Oxford University Press, forthcoming), in a chapter entitled "Death and Later Goods."

12. Nagel, pp. 8 and 10.

A. What Death Deprives Us Of

What Nagel seems to be suggesting, in the passage quoted above, is that death is *always* bad for the person who dies relative to the possibilities for good that one could imagine his life containing had that life continued. What is unclear, however, is why one should think that a set of possibilities that are possible only in the sense of being imaginable constitutes a relevant alternative to death for purposes of comparison. Simply to point out that there is an imaginable possible future life that a person might have had if he had not died seems insufficient to show that he met with a bad end. For we can also imagine possible future lives that the person might have had which would not have been worth living, relative to which his death could be judged not to be bad, or even to be good. Since desirable future lives and undesirable future lives are all equally imaginable, there seems to be no more reason to judge the person's death to be bad than there is to judge it to be good. But unless there is some nonarbitrary way of selecting, from among the many imaginable lives the person might have had, the one which can be considered the relevant alternative to death for purposes of comparison, then Nagel's focus on what would have been imaginable in the absence of death provides no basis for the evaluation of an individual's death at all.

What is relevant in evaluating the badness of an individual's death is not what might have happened if he had not died, but what would in fact have happened. When we evaluate death relative to what would have happened in its absence, we find that it is clearly not the case that death is always infinitely bad because it deprives its victim of an "indefinitely extensive" set of possible goods. Rather, the possibilities for good of which a person is deprived by death are limited by the fact that, had he not died when and how he did, he would have been condemned by his biology and circumstances to die within a certain limited period of time thereafter. The relevant alternative to death for purposes of comparison is not continuing to live indefinitely, or forever, but living on for a limited period of time and then dying of some other cause. So, other things being equal, we measure the badness of death in terms of the quantity and quality of life that the victim would have enjoyed had he not died when and how he did. (I will later explain how other things might be relevantly unequal.)

This approach to determining the badness of death supports the common view that it is in general worse to die earlier rather than later.[13] As a rule, a person who dies at thirty would, had he not died, have en-

13. As his text makes clear, Nagel himself believes that it is normally worse to die earlier rather than later. It is not clear, however, how he thinks this can be rendered consistent with his belief that even those who die of extreme old age thereby suffer a loss of indefinite possibilities for good. In Sec. II.A.4., I will suggest a way of reconciling his various intuitions.

joyed more goods before meeting with death from some other cause than a person who dies at eighty. This approach also supports the view that most of us take in cases in which it seems that death is not a bad thing for the person who dies, or is even a good thing. These are cases in which we feel that suicide would be rational, or euthanasia justified. The reason why death is not bad in these cases is that, if the person were not to die, the life that he would subsequently have would not be worth living. If, however, we assume with Nagel that death should be evaluated relative to the possibilities for good that would be imaginable in its absence, then it seems that we should regard death as an evil even in these cases. This makes it difficult to see how those who take Nagel's view can find cases in which suicide would be rational or in which euthanasia would even be conceptually possible.

Let us call the account of the badness of death which I have sketched the "revised possible goods account." In order to be able to explain our intuitive beliefs about the comparative badness of different possible deaths, this account requires further refinement. The necessary additions and qualifications will be developed in the following five subsections.

1. *Counterfactual Conditionals and the Problem of Specifying the Antecedent.*—One problem which the possible goods account faces can best be introduced by means of an example. Consider the case of a thirty-year-old man who died today of cancer. Call him Mort. According to the possible goods account, our evaluation of Mort's death must be based on a counterfactual conditional claim to the effect that "such-and-such would have happened if. . . ." The natural candidate for the antecedent of the counterfactual is, of course, "if he had not died." But this is in fact hopelessly vague, for there are countless ways in which he might not have died. To imagine that his death did not occur we must imagine that its cause did not occur, or did not lead to its expected effect. Yet there are not only various ways of understanding what the cause of his death was, but there are also, for each way of conceiving of the cause of death, various ways in which it would have been possible for the cause not to have operated. Suppose, for example, that we say that the cancer was the cause of his death. Then to suppose that he might not have died is to suppose that he might have been cured of the cancer, or that he might never have been stricken with it in the first place, or that he might have lived on with it in a nonfatal form. In short, we must suppose that he might not have died from the cancer, but there are various ways in which this might have happened.

Suppose, on the other hand, that we want to be more specific in identifying the cause of Mort's death. We might then say that the cause of his death was the immediate mechanism by which his death was brought about—a hemorrhage, for example. (This is what a pathologist who knew of Mort's condition would want to know if he were to ask what the cause of death was.) If this is how we conceive of the cause of death, then to

hypothesize that he might not have died is to suppose that *this* cause of death might not have operated.

Which of these various ways of spelling out the antecedent ("if he had not died at *t*") is the one on which we should base our evaluation of Mort's death? Certainly our choice of interpretation will make a difference to what our evaluation will be. For our estimation of what would have happened to Mort had he not died will depend on how we understand the circumstances of his not dying. If not dying would have involved being cured of the cancer, then it may be true that had he not died, he would have lived a long and prosperous life. Relative to that understanding of the antecedent, then, Mort's death can be regarded as a terrible tragedy. But suppose that his not dying is understood in such a way that it would have involved only the absence of the immediate mechanism of death (the hemorrhage). In that case we may suppose that, had he not died when he did, he would have lived on for only a short period of time until his death would have been brought about by some alternative mechanism associated with the cancer. Relative to this understanding of the antecedent, Mort's death was not tragic, and it may even have been a good thing (if, e.g., the brief future he would otherwise have had would have been filled with suffering).

Mort's case is in a way paradoxical, for there is both a clear sense in which it is true that his death today was not tragic—indeed, was perhaps on balance a good thing—as well as a clear sense in which his death was tragic. We can see now that each of these claims is true relative to certain ways of specifying the antecedent and false relative to others. His death today was perhaps a good thing relative to living on for a few days in agony, but it was bad relative to continued life free from cancer.

It seems unsatisfactory, however, to be left with a multiplicity of different and superficially conflicting evaluations of his death. What we want is a single, general, context-independent evaluation which tells us whether, all things considered, his death was good or bad, period. Our problem, then, is to find a method for picking out the single most general, context-independent way of understanding the idea that he might not have died. Let us call this *the problem of specifying the antecedent,* since what we are trying to do, in effect, is to assign precise and determinate content to the phrase "if he had not died . . . " in such a way that the completed conditional, if true, will provide the most general and context-independent assessment of what Mort has been deprived of, or has been spared, by his death.

How might we discover a method of specifying the antecedent in this and all other cases in which we seek a single overall evaluation of a person's death? I see no alternative but to propose various methods and test them against our intuitions by exploring their implications for a range of cases. Before proposing what seems a promising candidate, I will introduce some further cases which will not only help to illuminate

the structure of the problem but will also serve as useful tests for assessing the plausibility of proposed methods.

2. *The Problem of Causal Overdetermination.*—There is a familiar problem in ethics, which is that it is often difficult to know how to evaluate an act which causes injury if, had the act not been done, some other cause would have operated around the same time to produce a relevantly similar injury to the same victim. Following Feinberg, I will say that these are cases in which injury to the victim is causally overdetermined.[14] A similar problem arises for the evaluation of death when death is causally overdetermined in this sense.

Consider the following case. Joe is twenty-nine and a half years old. Schmoe has just turned thirty. Both are run over by a bus as they step off the curb. Our initial reaction is to think of both deaths as terribly tragic, Joe's being perhaps slightly more tragic than Schmoe's because of his slightly younger age. But suppose that, while Schmoe was in robust good health, it is discovered during the autopsy that Joe had a silent, symptomless, but invariably fatal disease that would have killed him within two months had he not been mown down by the bus. Our response to the discovery of this fact is to revise our initial assessment of the badness of Joe's death. Joe's death now seems considerably less bad than it would have been had he not had the condition since, given the fact that he had the condition, all he lost in being hit by the bus was at most two months of further life. Our revised response is thus to think that Schmoe's death was the more tragic of the two, other things being equal, even though he was older than Joe.

The view that Joe lost only relatively little by dying in the accident is supported by the fact that it explains certain intuitions. Suppose that one knew about Joe's disease and that one was in a position to snatch either Joe or Schmoe, but not both, out of the path of the bus. It seems that it would be better to save Schmoe rather than Joe if other things, such as their importance to other people, were equal. The reason why it would be better to save Schmoe is simply that he would have more to lose by dying than Joe would.

This may seem straightforward. In fact it is not. When we learn the results of the autopsy, our inclination is to think that Joe's death was less bad than it initially seemed. But if we then report the results to Joe's grieving mother, it is very unlikely that she will find any grounds for consolation in the fact that he would soon have died anyway. She might reason as follows. Suppose that Joe had not been killed by the bus. Then he would have died sometime within the ensuing two months from the disease. What would we then have said of *that* death? If we were to ask

14. Joel Feinberg, "Wrongful Life and the Counterfactual Element in Harming," *Social Philosophy and Policy* 4, no. 1 (1986): 145–78. Normally the phrase "causal overdetermination" is reserved for cases in which two or more individually sufficient causes operate simultaneously to bring about some event.

what would have happened had he not then died from the disease, the answer could well be that he could have been expected to enjoy a long and fruitful life. In that case we would have said that his death from the disease was a terrible tragedy—a case of a young man cut down in his prime. But if death at so early an age from the disease would have been tragic, how can it be *less* bad when he is killed today by a bus at an even earlier age?

I will return to the question of Joe's misfortune in Section II.A.4. I want now to introduce another case—the case of a young officer in the cavalry who was killed in the charge of the Light Brigade. This officer was among the leaders of the charge and was shot quite early by a soldier named Ivan. Suppose that, had he not been shot by Ivan, he would have been killed within a few seconds by a bullet fired by Boris, who also had him within his sights. Our natural response to this case is to say that the officer's death was a grave misfortune, depriving him of many years of life. Yet it would seem that this case is like that of Joe in being a case in which death within a relatively short span of time is causally overdetermined. Thus, if we accept that Joe lost at most two months in being hit by the bus, should we not also conclude that in this case all the officer lost in being shot by Ivan was a few seconds of life, so that his death was hardly a misfortune at all?

Obviously we cannot accept this. Yet in the case of Joe it is *not* implausible to regard his death in the accident as considerably less bad than it would have been had there been no other cause of death which would soon have killed him if the bus had not. Is there some difference between the cases which explains our differing reactions to them? I think that there is and that examining that difference will point us toward a plausible resolution of the problem of specifying the antecedent.

3. *A Proposal.*—The solution I will offer to the problem of specifying the antecedent is based on a variant of the general analysis of the truth-value of counterfactuals associated with the work of David Lewis and Robert Stalnaker.[15] What the variants of this form of analysis have in common is the view that a counterfactual conditional is true if and only if the consequent is true in the possible world (or in all of the possible worlds) in which the antecedent is realized that is (or are) closest to the actual world, where closeness is determined by the application of some similarity metric. I will simply assume that some theory of this sort is true and that the version that I appeal to is a reasonable candidate for the best version. I will not try to defend or even to make explicit all the details of this version.[16] Indeed, as will soon become apparent, the proposal I will advance raises far more problems of metaphysics than there is

15. See Robert C. Stalnaker, "A Theory of Conditionals," *American Philosophical Quarterly*, monograph no. 2 (1968); and David Lewis, *Counterfactuals* (Princeton, N.J.: Princeton University Press, 1973).

16. I suspect, though I cannot prove this, that my proposal could be reformulated without significant loss in terms of other variants of the general theory of counterfactuals.

space to address here, so one should regard what follows as nothing more than a tentative sketch of a solution.

It is perhaps tempting to think that the solution to the problem of specifying the antecedent is given more or less directly by the Lewis-Stalnaker theory. One might think, for example, that we can identify the appropriate antecedent-world as the possible world in which the person does not die but which otherwise deviates minimally from the actual world (at least up to the time at which the person died in the actual world). This, however, would be a mistake. Return to the case of the cavalry officer. Suppose (what is not implausible) that the closest possible world in which the officer does not die from Ivan's bullet is one in which he is grazed by Ivan's bullet and then killed by Boris's bullet. In that case our answer to the question of what would have happened had the officer not died when and how he did will be that he would have lived for a few seconds, and then he would have been killed. This leads to the unacceptable conclusion that his actual death was hardly a misfortune at all.

In the case of the cavalry officer, as in the case of Mort, what we believe about what would have happened had he not died when and how he did depends on how we specify the cause of his death. If we single out the shot fired by Ivan as the cause and then ask what would have happened had that cause not operated (e.g., because Ivan's hand shook, or his gun jammed), we get the answer that the officer would soon have been killed by Boris. We might, however, identify the cause of the officer's death differently—for example, as his being shot in the charge. If we ask what would have happened had he not been shot in the charge, the answer may well be that he would subsequently have led a long and happy life. Of course, as in the case of Mort, these two claims about what would have happened lead to radically divergent evaluations of the officer's death.

The problem of specifying the antecedent, then, is a problem about identifying and delimiting the cause of death. In asking what would have happened had a certain person not died, we are asking what would have been the case had the cause of his death not operated. The many different ways of specifying the antecedent thus correspond to different ways in which the chain of causation leading to the person's death might have gone off course or been disrupted—different ways, in short, in which various individually or jointly necessary conditions of the person's death might not have occurred. The problem of specifying the antecedent, therefore, is the problem of determining which elements from the totality of causally relevant factors in the etiology of death we should imagine not occurring when, in order to obtain a single overall evaluation of the death, we imagine what would have been the case had the victim not died.

Return now to the comparison between the case of the officer and the case of Joe. An important difference between the two cases would

seem to be that, while the two potential causes of the officer's death (the threat from Ivan and the threat from Boris) both seem to be parts of the same causal sequence, the two potential causes of Joe's death (the bus accident and the fatal condition) are not. Hence in the possible world in which the entire causal sequence containing the immediate cause of the officer's death is absent, there are no other causes of death lurking in the foreseeable future. But if in Joe's case we subtract the causal sequence leading to his death in the bus accident, we do not thereby remove the other cause of death—namely, the fatal condition.

What this suggests is that we can secure the intuitively correct counterfactual claims in these cases if our formula for specifying the antecedent is to subtract the entire causal sequence of which the immediate cause of death is a part. Thus, if the entire causal sequence leading to the officer's being shot by Ivan had not occurred, then he would presumably have gone on to lead a long and prosperous life; whereas if the entire causal sequence leading to Joe's being hit by the bus had not occurred, he would still have died within two months of the condition. Because of this, we may regard Joe's death as less tragic than that of the officer.

This formula, while promising, requires considerable tightening up. One problem, for example, is to give sufficient content to the notion of a causal sequence to allow for the individuation of causal sequences in the way that the formula requires. Here is one suggestion for dealing with this problem. Let us define the *transitive cause* of an event E as follows. If C is the immediate or proximate cause of E, then the transitive cause of E is the set of all the events that form part of the chain of causes leading to C. This set is understood to be both complete and closed, in the sense that all and only the members of the set satisfy the following condition: if Ci is a member of the set and Cj is a cause of Ci, then Cj is a member of the set. Given the notion of transitive cause, we can now state more precisely what it means to imagine a possible world in which the entire causal sequence containing a certain death is absent. To imagine such a world is simply to imagine a world which is like the actual world except that the entire transitive cause of the death is absent, along with those events for whose occurrence some element or elements of the transitive cause were a necessary condition given the laws of causation that hold in the actual world. So, for example, if we imagine that the transitive cause of the officer's being shot by Ivan did not occur, we must presumably imagine that the Crimean War did not occur, in which case the threat from Boris would not have occurred either.

It is important to notice that the plausibility of this suggestion depends on our distinguishing between the cause of some event and the causally relevant conditions of that event. The transitive cause of E, as I understand it, consists of a chain of causes (namely, the cause of the cause of the cause . . . of E) and not the set of all the causally relevant conditions of E. For example, the transitive cause of the officer's death presumably includes the occurrence of the battle and thus the occurrence of the war,

but not the event of Ivan's birth, or the presence of oxygen in the air for Ivan to breathe, or even the fact that the officer was not wounded and sent home just prior to the battle—though all of these are causally necessary conditions of the officer's actual death. If the transitive cause of the officer's death were to include these and all the other causally relevant conditions of his death, the proposal I am putting forward would obviously be hopelessly implausible.[17]

Having sketched certain theories, concepts, and distinctions on which my proposed solution to the problem of specifying the antecedent is based, I am now in a position to state the proposal in full. Let t be the time at which some person died. Our overall, objective evaluation of how bad or good his death was for him will be based on a counterfactual claim about what would have happened to him if he had not died at t. Let the antecedent of the relevant counterfactual be "if the entire transitive cause of his death had not occurred. . . ." To complete the counterfactual, we consult the possible world in which the antecedent is realized which is closest to the actual world up to t.[18] In assessing comparative similarity, we give nomological similarity lexical priority over factual similarity. That is, we hold the laws of causation constant across possible worlds.[19] Then we simply let the future unfold in this world in accordance with the laws of causation that hold in the actual world, and see how the person fares. If, for example, in the closest world in which the antecedent is realized the person goes on to live a long and happy life subsequent to t, then it is true that if he had not died he would have lived a long and happy life. His death was then bad in rough proportion to the amount of good that his life would have contained. (If the future is causally underdetermined by the present, then he may fare differently in different possible futures which are all compatible with running the closest antecedent-world forward in accordance with the laws of causation. In that case, the sum of the goods that would have been causally possible for the person will presumably exceed the amount of good that would have been contained in any single future life. Thus, in attempting to calculate the losses that the person suffers through death, we should presumably weight each good that would have been causally possible for him according to how

17. I have no analysis of the distinction between cause and causal condition to offer, nor any view about whether the distinction marks a real difference or is simply context dependent. I here rely on our intuitive sense of what counts as a cause and what counts as a causal condition.

18. Lewis contends that there may be more than one closest possible world in which the antecedent is realized. I must ignore this complication here.

19. Unless, perhaps, we are forced to consider causally impossible worlds by the fact that the antecedent itself is causally impossible. (More on this later.) In insisting that we alter matters of fact before we alter the causal structure of the actual world, I am assuming a version of the similarity theory that many writers, including Lewis, would reject. For arguments in favor of banning miracles in closest antecedent-worlds, see Jonathan Bennett, "Counterfactuals and Temporal Direction," *Philosophical Review* 93 (1984): 57–91, esp. sec. 7.

probable its occurrence would have been had he lived. Again, however, there is no space to pursue this further complexity.)

This proposal provides what seem to be the intuitively correct answers in the cases we have considered so far. For example, it implies that the cavalry officer's death was tragic since, if the transitive cause of his death had never occurred, he would not have been threatened by either Ivan or Boris and so, we may assume, would have gone on to live a long and happy life. It also implies that Joe's death was significantly less bad than Schmoe's since, in the closest possible world in which the transitive cause of his death is absent, Joe would have lived on for at most two months; whereas, in Schmoe's case, removing the transitive cause of death would have left him with many years of life remaining. Finally, the proposal implies that Mort's death was a tragedy comparable to those suffered by the officer and by Schmoe, since the relevant counterfactual situation with which we compare his death is one in which he never contracts cancer in the first place.

It should be emphasized that these judgments represent only our most general, maximally context-independent evaluations of these deaths. As I noted earlier in introducing the problem of specifying the antecedent, there are many other possible evaluations of each of these deaths which are based on other imagined ways in which the causal chain leading to the death might have been interrupted so that the death would not have occurred. These alternative evaluations need not be mistaken. On the contrary, they can be quite important for certain purposes, such as guiding action at moments when, though it is not possible to prevent the entire transitive cause of some potential death, it nevertheless is possible to intervene in the causal chain in some other way. For example, the doctors may wish to know whether it would be in Mort's interests for them to avert the hemorrhage which they predict will occur unless he is given a certain medication. Whether it would be in or against his interests depends at least in part on whether his death from the hemorrhage would be good or bad for him at that time (i.e., relative to what would be in store for him if, in his present condition, his death from the hemorrhage were to be prevented). In making this decision, the doctors should be guided by their assessment of what Mort's life would be like in the closest possible world in which the causal chain which will otherwise lead to his death is broken or diverted at the point at which it is possible for them to intervene.

So evaluations of this sort need not be wrong, and indeed are indispensable in a variety of decision-making contexts. But, unlike the evaluations that issue from the application of the procedure outlined above, they are essentially context dependent. They can help to guide our action in relevant circumstances, but they do not tell us, for example, to what extent it is appropriate to feel pity for someone because he has died.

4. *Deprivation of Future Goods and the Global Evaluation of Lives.*—I have claimed that the proposal sketched above gives the right answer in

the case of Joe, but earlier I raised doubts about whether it is really our intuition that his death was less tragic than that of Schmoe. Even when we discover that Joe suffered from the fatal condition, we, along with his mother, retain our sense that he has suffered a terrible misfortune. And in fact we are right. Our mistake is to identify this misfortune with the event of his death. For it remains true that the event of his death was not itself a grave misfortune for him, since at most it deprived him of only two months of life. Rather, his misfortune consists in his having been deprived of a future in which there would be possibilities for good. It is because the absence of future possibilities for good (at least of an active or experiential kind) was in this case overdetermined, in that it was guaranteed by the disease quite independently of the bus accident, that the misfortune is not ascribable to his death alone. It is instead ascribable to the fact that he was killed by the bus *together with* the fact that if he had not been killed by the bus he would have died soon of the disease. It is this which is the proper object of his mother's great grief.

It may, however, be objected to this that the misfortune which I claim that Joe suffered is one that befalls us all, so that in the end we are all equally unfortunate, equally deprived. Consider, for example, the case of a person who dies from extreme old age, at the biological limits of human life. Call this person Gerry. It seems that Gerry's case is, according to my view, relevantly like Joe's. His death within a short span of time is overdetermined. Suppose that he died of a heart attack. Even if the transitive cause of the heart attack had not occurred, Gerry would nevertheless have died soon of some other cause associated with the fragility of old age.[20] For this reason most of us readily concede that his death was not a great misfortune for him; we acknowledge, in this case, that very few goods would have been causally possible for him had he not died. But, an objector might claim, the fact that he would have had little to look forward to had he not died is, according to my view, itself a great misfortune—precisely the same misfortune which I have claimed that Joe suffered. But if Joe and Gerry both suffer this misfortune equally, and if neither one's death is particularly tragic, then it would seem that the problem raised in the passage from Nagel quoted in Section II—namely, that we all suffer infinite losses—has reemerged in a slightly altered form. For it now seems that, on my view, we all *do* suffer infinite losses, though these losses are not attributable to death alone. But if this is right, how can we explain our feeling that Joe has suffered a greater misfortune than Gerry and is more deserving of our pity?

One reply to this challenge might appeal to the idea that the failure to realize some good is less bad the less causally possible it was to realize

20. I assume that we may rule out counterfactual claims based on antecedent-worlds in which the aging process is arrested or reversed on the ground that any such world would be insufficiently similar to the actual world in terms of any plausible similarity metric to provide truth-conditions for a counterfactual.

that good.[21] If that were true, then Joe would be the more unfortunate of the two because future goods were causally more possible for him than for Gerry, for whom the absence of future goods was multiply overdetermined by a variety of potential causes of death as well as by aging, disease, and so forth. But in fact this hardly seems to matter. Even if the absence of future possibilities for good were equally overdetermined in the two cases (e.g., if there were twenty Murder Inc. hit men independently trying to kill Joe when he was hit by the bus), we would still grieve more for him and think him more unfortunate than Gerry.

There is a better reply. This reply begins by embracing the allegedly absurd view that Gerry suffers a serious misfortune. Indeed, it is compatible with the view that ultimately we *all* suffer a great misfortune—even those of us who live the longest, richest lives. This is not the best of all possible worlds. We are all subject to aging beyond maturity, disease, injury, death, and so on. Were it not for these various evils, each of us would enjoy the prospect of an indefinitely extensive succession of possible goods. Perhaps it is this fact that Nagel has in mind in the quotation cited earlier. If so, his mistake is to think that the deprivation of future possibilities for good that we all suffer is attributable to death when it is in fact overdetermined by a package of evils of which death is only one.

The fact that we all suffer this deprivation equally does not, however, imply that we are all equally unfortunate. Consider again the comparison between Joe and Gerry. Even if their losses are ultimately the same, their gains from life have not been. The explanation of why there is less reason to grieve for Gerry is simply that he has had a fair share of life. Relative to reasonable expectations, he has had a rich and full life. Joe's life, by contrast, has fallen short of what he could reasonably have expected to gain from life.

Even though those who gain more from life than most of us *do* suffer a misfortune when they die or when their lives become such that death, in the circumstances, would be good for them, they are nevertheless seldom to be pitied. Feelings such as pity have to be adjusted to normal or reasonable expectations. Consider the case of someone who badly wants to travel to distant parts of the universe. It would be a mistake to suppose that he suffers no misfortune in being unable to do what he most wants to do (something which we would all agree to be well worth doing if it were possible—i.e., his desire, though idle, is not frivolous). But the fact that his desire is frustrated does not evoke pity. This is because pity is appropriate only in cases in which a person is unable to fulfill reasonable expectations given the circumstances of human life.[22]

21. Compare Nagel, "Death," p. 9, where he appears to suggest that whether (and perhaps to what extent) the failure to realize some future good counts as a misfortune depends on how possible the relevant good is.

22. Our evaluation of a person's global good or ill fortune should thus be relativized to our perception of the norm for a certain sort of human life. This should not be confused with the common but nevertheless unacceptable idea that we should evaluate a person's

If we assume that there is a limit to the amount of good that any life can contain, then it is natural to conclude that there is a correlation between what a person gains from life and what he loses when possibilities for good cease to be available to him. But, when we assess our lives relative to the possibility of a desirable immortality (which is what we do when we measure the misfortune inflicted on us jointly by aging, disease, physical vulnerability, death, etc.), it ceases to be true that, if other things (such as the quality of life) are equal, then the more a person gains from life, the fewer his losses will be when deprived of future possibilities for good. Since this correlation breaks down in this case, it is not unreasonable to measure the comparative misfortune that a person suffers in being deprived of future possibilities for good in terms of the extent to which the deprivation limits his gains from life rather than in terms of the extent to which it increases his losses.

There is, however, an objection to this claim. Consider the case of two elderly men, Faust and his friend Fred. Both have had lives of roughly equal length and quality. Faust is approached by Mephistopheles and given the elixir of youth—in this case with no penalty to be paid later on. The elixir gives Faust the body of an eighteen-year-old. He then sets out joyously on a new life. But on the first day of this new life he and Fred are both run over by a bus. What should we say about this case? Both men suffer the misfortune of being deprived of future possibilities for good. Moreover, the limitation on their gains from life is the same in each case. Neither has gained more from life than the other. But surely Faust suffers a greater misfortune than Fred?

In fact Faust is no more unfortunate than Fred *in global terms*. It is true that Faust is killed just as he is about to embark on a second youth. What this means is that Faust's *death* is more tragic than Fred's. The event of his being hit by the bus is worse for Faust than being hit by the bus is for Fred. Thus if both Faust and Fred have been mortally injured by the bus and lie dying in their hospital beds, Faust will naturally feel greater bitterness about being hit by the bus than Fred will. It is often harder to bear the loss of some recently acquired good than it is never to acquire the good in the first place. But never to acquire some good is also a misfortune. So, just as Faust is unfortunate in having the benefits of the elixir snatched away from him, Fred is also unfortunate in never having been given the elixir.

I will now conclude this section by briefly considering one final problem case—a case which seems to pose a problem for any account of the badness of death. Suppose that there is a young woman who dies of a genetic condition that is both strongly incurable (in the sense that it is causally impossible to suppress its effects) and essential to her identity

---

death relative to the possibility of his living out a lifespan that is normal for persons in his society. Among the various objections to the latter idea is the fact that it seems to entail that death cannot be bad for those who have already lived longer than the normal lifespan.

(in that anyone born without the condition would not have been *her*). Call this woman Genette. If we ask what would have happened if the transitive cause of her death had not occurred, the answer seems to be that she would never have existed. The possible world in which the transitive cause is absent is therefore not one in which she fares better than she does in the actual world. So it may seem that we are forced to conclude that she suffers no misfortune in dying of the condition, nor indeed in having the condition. Yet many of us share Nagel's intuition that death can be a misfortune even when it is strongly impossible that it should not occur.

I will not defend a particular solution to this problem, but will simply outline three possible responses, all of which have a certain plausibility. The first response is just to accept that Genette's death is not a misfortune on the ground that she has had the longest life that she could have. The second response concedes that there are no possible worlds in which she lives longer but rejects the inference that her death was not bad for her. This response appeals to counterpart theory. It holds that the truth-conditions for claims about what would have happened had she not had the condition are provided by closest possible worlds containing the closest counterpart to Genette in whom the genetic defect is absent. The fact that this counterpart would not, strictly speaking, be *Genette* would not necessarily undermine the relevance of events in these worlds to evaluative claims about Genette's life. According to some versions of counterpart theory, such as Lewis's, a counterpart in some possible world is never identical with the original object in the actual world.

The third response involves two deviations from the proposal advanced earlier in Section II.A.3. The first is that we should ask, not what would have happened if the transitive cause of Genette's death had not occurred, but instead what would have happened if the effects of her condition had been suppressed. The second is that we should allow that the closest possible world in which that antecedent is realized may contain some violation of the laws of causation that hold in the actual world. We may perhaps plead necessity in both cases. For example, the justification for the second deviation might be that, while we must preserve the laws of causation wherever possible, we are simply forced to admit closest possible worlds containing causal impossibilities when we consider antecedents which are themselves causally impossible. Consider for example, the claim that, if the law of universal gravitation were an inverse cube law rather than an inverse square law, objects would then attract one another less strongly at great distances and more strongly at short distances. This claim seems true, though the possible world which provides its truth-conditions is causally impossible. Notice, furthermore, that unless we make this concession we will be unable to sustain my earlier claim that we all suffer a misfortune in being physically vulnerable, susceptible to disease, and so on. If we accept that we are unfortunate in being vulnerable to injury, on the ground that if we were not, our lives would be both

longer and happier, then we should be able to make the analogous claim in Genette's case that she is unfortunate in suffering certain effects of her condition, among which we include the fact that it leads to premature death.

B. Death and the Degree of Psychological Connectedness

Let us turn now to one final problem for the possible goods account. In general our belief is that death is worse the earlier it occurs. This seems to be implied by the possible goods account. For it would seem that, the earlier death occurs, the greater are the losses which the victim suffers and, correspondingly, the fewer are the gains which he derives from life. But many people feel that there are instances in which an earlier death is not worse for the victim, even if the whole of the life that is lost would have been well worth living. They believe that death at, say, one month after conception, or one month after birth, is not worse, or more tragic, than death at, say, twenty, or forty. Rather, death at twenty, or forty, is in fact *worse* than death before birth, or perhaps even in early infancy. How can this view be explained?

I think it is best explained by appealing to the theory which holds that the criterion of personal identity over time is psychological continuity.[23] According to this theory, continuing to exist as the same person over time essentially involves the holding of certain psychological relations—such as that between an experience and the memory of it, or that between an intention and the later act that carries it out. This presupposes that it is a necessary condition of being a person *at all* that an entity should have at least some of the psychological states or characteristics that are the ingredients of these various relations. The theory is, moreover, reductionist. It holds that the existence of a person is ultimately reducible to the existence of a certain set of psychological events and the relations among them.

Since there is no mental life at all associated with a human fetus in the early stages of gestation, the fetus cannot be a person during this period. If, as I believe, personhood is an essential property, it follows that the fetus during this period cannot be the same individual as the person who would later exist if the fetus were to follow its natural path of development. Thus if a fetus dies early in its career, *it* suffers no loss at all. Since it has no mental life, its death cannot involve the loss of anything that is of value to it at that time; and the loss of the future life that its death involves is also no loss to it, since that life would not have

23. This theory is most persuasively developed in Derek Parfit's *Reasons and Persons* (Oxford: Oxford University Press, 1984), pt. 3. For the purposes of this paper I will assume that this theory is true. It is not, however, strictly necessary for my argument that we should accept this theory. While the claims that are essential to my argument—e.g., that it can be rational for a person to discount for diminished psychological connectedness—are probably best supported by this theory, they are in principle compatible with certain other theories. Hence one is not necessarily committed by the acceptance of my argument to any particular theory of personal identity.

been *its* life but would instead have belonged to the person.[24] Thus the reason why death at this early stage is not worse for the victim than death at age twenty is that death at this stage is not bad for the victim at all.

Later in the course of the development of a human fetus, when the rudiments of a mental life begin to appear, a person begins to develop. But the psychological attributes which are the constituent elements of personhood do not appear all at once, but instead appear gradually, so that the mental life that develops in association with the human organism becomes richer, more sophisticated, and more unified as the organism matures and develops. Since the existence of a person is nothing more than the existence of certain psychological states and their interrelations, and since the appearance of the states and relations that are constitutive of a person is a gradual process, the development or coming-into-existence of a person is also gradual.[25] It seems to follow that the existence of a person can be a matter of degree. During the period when the mental life associated with the developing organism remains relatively primitive, the person does not yet fully exist. What exists is only a potential or developing person—an entity in the process of becoming a person. Only when this sequence of mental activity reaches a certain level of complexity and sophistication does the person fully exist. There is, however, no determinate threshold of complexity and sophistication such that in passing the threshold the person suddenly achieves full existence. Rather, there is a period during which, while what exists is clearly at least a potential or developing person, it is indeterminate whether it has yet fully become a person.[26]

Let us suppose that even one month after the birth of the human organism the person is not yet fully realized. The one-month-old infant is still only a potential, or developing, person. If the infant dies, its losses will be less than those that a fully developed person normally suffers

24. It might be objected that, since my present physical organism is certainly the same organism as the fetus with which it is physically continuous, the claim that *I* am not identical with that fetus implies that I am not my physical organism. I accept this objection. Since I could cease to exist even if my organism were to continue to exist and indeed continue to live, it would seem that I cannot be identical with my organism. This, however, may be compatible with the claim that there is a further sense in which I am my organism. To borrow an old analogy, the sense in which I am my organism is the same sense in which a statue is the clay of which it is made. Just as the statue is constituted by the clay without being identical with it, so I consist of or am constituted by my organism without being identical with it. (These issues are ably discussed by W. R. Carter in "Do Zygotes Become People?" *Mind* 91 [1982]: 77–95, and in "Once and Future People," *American Philosophical Quarterly* 17 [1980]: 61–66; and by Warren Quinn in "Abortion: Identity and Loss," *Philosophy and Public Affairs* 13 [1984]: 24–54, though both of these writers defend a different position from that advocated here. I discuss the issue of the relation between persons and their organisms, as well as the other issues raised in this section, in considerable detail in chap. 2 of *The Ethics of Killing*.)

25. For a cogent defense of the view that persons, even if they are substances, come into existence gradually, see Quinn, pp. 33–40.

26. Compare Parfit, sec. 79.

through death. The future life that is lost when the infant dies would have been the life of a fully existing person, and hence it belonged only partially to the infant itself. This helps to explain why death in early infancy is not worse for the victim than, say, death at twenty. It is only after a person becomes fully real that death is normally worse for him the earlier in his life it occurs.

A different though closely related way of supporting the same conclusion appeals to a reductionist view about what matters in personal survival. Most of those who accept the psychological continuity theory of personal identity believe that identity is not what matters in personal survival. What matters is not that there should later exist someone who is *me* but instead that there should later be someone with whom I am sufficiently closely psychologically connected. Psychological connectedness, in other words, provides the principal basis for egoistic concern about the future.[27] Psychological connectedness, however, is a matter of degree. This suggests that the weaker the psychological connections between a person now and the same person later, the weaker his grounds will be now for egoistic concern about his later life.

Even if we suppose—what seems doubtful—that a one-month-old human being can be regarded as a fully real person, there will be few, if any, direct psychological connections between the person at one month and the person later in life. Hence our grounds for concern about the infant's future life for *its* sake will be correspondingly weak. If the infant dies, there is a sense in which *its* losses are less than those that an adult human being normally suffers through death, since the future it loses would have been less closely connected to it in the ways that provide grounds for egoistic concern—or perhaps, in this case, grounds for concern for its sake. (Of course, the infant's death may mean the loss of a glorious future life. But the only loss that matters significantly where the infant—qua infant—is concerned is the loss of those parts of its future life with which it would have been psychologically connected.)

C. The Revised Possible Goods Account

What this appeal to the theory of personal identity suggests is that the possible goods account, to be plausible, cannot assess the losses involved in death by simply summing up the goods that life would likely have contained were it not for death. These goods must instead be weighted in such a way that the loss through death of some future good will count

27. Ibid., chap. 12, esp. secs. 89 and 90. While psychological connectedness is what matters most, a simple argument suggests that psychological continuity also matters. Suppose that I am now psychologically connected with myself at t_1 but not with myself at t_2. Suppose that, because of the lack of any connections, I do not now have any direct concern about what happens to me at t_2. I nevertheless realize that at t_1 I will care about what happens to me at t_2, since at t_1 I will be connected to myself at t_2. Thus, since I now care about my interests at t_1, I now have an indirect reason for caring what happens to me at t_2, since what happens to me at t_2 will matter to me at t_1.

for more the closer the psychological connections would have been between the person as she was at the time of her death and as she would have been at the time at which she would have received or experienced the good.[28]

There are other ways in which the possible goods account can be revised or extended in order to be better harmonized with our common conception of the badness of death. It should, for example, make some provision for the importance of desire. While it seems implausible to suppose that one can fully account for the badness of death in terms of the fact that death frustrates so many of the victim's desires, it does seem to make a difference whether and to what extent the possible life which death prevents would have or might have contained experiences or activities that the victim actually cared about during his life.[29] So, for example, even if the possible life that is lost through death would have contained a rich variety of goods, and even if the victim would have cared about and valued these goods had he lived to experience them, his death may seem less tragic if these goods would have been ones that he cared little or nothing about while he was alive.

This might be partially explained in terms of the assumption that the change in values that would have been required for the person to come to appreciate the goods would have involved a weakening of psychological connectedness. But this is not a complete explanation. The desires and concerns a person has are important in assessing the badness of her death independently of the fact that desires and values are constituent elements in psychological connectedness. It is an important part of the explanation of the badness of death that death frustrates the victim's desires, retroactively condemns to futility her efforts to fulfill them, and generally renders many of her strivings vain and pointless.[30] In particular,

28. It might be argued that these implications of the theory of personal identity themselves imply that personal growth, evolution, and improvement are, in prospect at least, undesirable because they would involve a weakening of psychological connectedness and hence reduce the value to oneself now of goods that one might acquire or experience in the future. Two points may be made in reply. First, personal growth and improvement need not involve any weakening of psychological connectedness, since changes in one's character can occur in fulfillment of one's earlier desires and intentions. In these cases the relevant changes actually constitute psychological connections which bind the later person more closely to his earlier self. Thus a life in which there is substantial change and diversity will, if the diverse elements are connected by threads of desire and intention, be unified in a deeper and richer way than a life in which one's character remains static. Second, even in cases in which personal improvement simply *happens* to a person, and so may involve some weakening of psychological connectedness, it may still be desirable, all things considered. For the weakening of connectedness may be outweighed by the greater importance of, e.g., an enhancement of virtue or rationality.

29. On the claim that the badness of death can be explained solely in terms of the frustration of desire, see McMahan, *The Ethics of Killing,* chap. 1, and the references to other work, e.g. by Bernard Williams, cited there.

30. This point has been well expressed by Michael Lockwood, who writes, "Set against an ideal of human life as a meaningful whole, we can see that premature death can, as it

death is worse to the extent that sacrifices have been made for the future, or to the extent that there has been an *investment* in the future.

Let me now review the results so far of our examination of the badness of death. Death is bad for a person (or developing person) at any point in his life, provided that the life that is thereby lost would on balance have been worth living. Other things being equal, the badness of death is proportional to the quality and quantity of the goods of which the victim is deprived. But, in assessing the goods of which the victim is deprived, we must weight them both according to the extent to which they were desired by the victim and according to how closely psychologically connected the person would have been at the time of acquiring them to himself at (and perhaps prior to) the time at which he in fact died. A short but perhaps helpful way of summarizing these results would be to say that death is worse the greater the potential unity and coherence of the life it disrupts.

It is perhaps worth emphasizing, in concluding this section, that, while the degree of unity and psychological connectedness in a life is relevant to assessing the value of the life or the extent to which the life is worth living, even a high degree of unity and connectedness is not a sufficient condition for a life's being well worth living. For not only can a life be highly unified and yet be of a very low quality, but it can also be highly unified around projects or aspirations that are trivial, absurd, contemptible, vile, or even evil. Thus the account I have sketched of the badness of death, and the implied account of the goodness of continued life, is merely formal. It is an account (and only a partial account at that) of the *structure* of the good life, and thus is compatible with different conceptions of the *content* of the good life. A more complete account would have to provide criteria for evaluating the meaningfulness of the different ways of living even highly unified lives.

III. A PARADOX

As we have seen, the view I have developed implies that death at, say, thirty-five is normally worse than death at one month. This seems plausible, and coincides with our intuitions. For, when a person dies at thirty-five, the life she loses is normally one with which she would have been closely psychologically connected and in which she would have enjoyed many of the goods that she valued, desired, and strived for. But neither of these points is true in the case of the life that is lost when a person dies

---

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-air, roads that lead nowhere" (Michael Lockwood, "Singer on Killing and the Preference for Life," *Inquiry* 22 [1979]: 167). This point, and the point about desire, reinforce the explanation developed in Sec. II.B of the view that death very early in life is not worse than death later in life.

at one month. It therefore seems plausible to say that a thirty-five-year-old normally stands to gain more from continuing to live than a one-month-old does, and that a thirty-five-year-old loses more through death. And this provides good grounds for thinking that death is worse at age thirty-five than it is at one month.

Let us suppose, however, that there is a newborn baby with a condition that, if untreated, will cause the baby to die at one month. The condition can be cured, but the treatment itself has the effect of making it inevitable that the person will die at the age of thirty-five. Should the baby's condition be treated? Most of us feel that, if the person's life could be expected to be worth living, then the baby ought to be treated. But this seems to indicate that we prefer for his sake that the person should die at thirty-five rather than at one month. But how can that be, given our earlier conclusion that death is worse at thirty-five than it is at one month?

We might say that to have more life that is worth living is always better than to have less, other things being equal. But that seems simply to restate the view that, if further life would be worth living, death is always worse the earlier it occurs. And that is what we have denied. We seem on the one hand to believe that death at thirty-five is worse, or more tragic, than death at one month. But now consideration of the case of the baby with the treatable condition leads us to believe that death *is* worse the earlier it occurs. We believe that it would be worse for this individual to die at one month than to die at thirty-five.

This paradox challenges the view about the badness of death that I have sketched because it seems to suggest that we feel compelled both to affirm and to deny the implications of that view. Is there a way in which the paradox might be dissolved? One suggestion might be that the paradox arises from our treating two different but superficially identical comparisons as if they were the same comparison. Consider, by way of analogy, the comparison between the following two choices. The first is a choice between death at a very early age—say at nine months after conception—and death at a much later age—say at thirty-five—for the *same person*. The second choice is between death at a very early age for one person and death at a much later age for *another* person. Both of these choices arise in a case in which a thirty-five-year-old woman who is nine months pregnant and will die unless a craniotomy is performed on the fetus. Suppose that, if the craniotomy is not performed, the baby will survive but, because of an inherited condition, will live to be only thirty-five. If we look just at the case of the fetus, we feel that it would clearly be worse for it to die now rather than at thirty-five. But if we compare the death of the mother now, at age thirty-five, with the death of the fetus now, most of us feel that the death of the mother would be the worse of the two.

While focusing on the case of the fetus alone might lead us to conclude that death at nine months after conception is worse than death at thirty-five, focusing on the comparison between the death of the fetus and the

death of the mother tempts us to conclude that death at thirty-five is worse than death at nine months after conception. But, it might be argued, there is here only a surface paradox. For the terms of the two comparisons between death at nine months after conception and death at thirty-five are not the same. When we focus on the case of the fetus alone, we are comparing death at nine months with death at thirty-five rather than earlier, at nine months. In this case, therefore, choosing death at thirty-five would mean that the person would *gain* thirty-five years of life. But when we compare the death of the fetus with the death of the mother, we are comparing death at nine months rather than at thirty-five with death at thirty-five rather than much later, at the end of a full life. Death at thirty-five for the mother, unlike death at thirty-five for the person in the previous comparison, involves no gains. Hence it is not surprising that we get different answers in the two comparisons. Death at nine months *is* worse than death at thirty-five rather than at nine months; but it is *not* necessarily worse than death at thirty-five rather than at some much later time.

It may seem that the paradox as I stated it earlier trades on exactly the sort of confusion I have uncovered here. On the other hand, even here when we consider the case of the fetus alone we conclude that it would be worse for that individual to die at nine months after conception than to die at thirty-five, and this alone seems inconsistent with our earlier claim that prenatal death is less bad for the victim than death at thirty-five. To dissolve the paradox completely it seems that we must recognize that the comparative badness of different possible deaths does not necessarily determine which death it would be worse to suffer. While it is usually worse for a person to suffer a worse death (e.g., to die at forty rather than at fifty), the case of the fetus suggests that it can sometimes be better to suffer a worse death. Similarly, it can also be worse to have a less bad death—for example, when death becomes less bad for a person because the life that he would otherwise have had has now become less good than it might have been. In the latter case, the fact that his death has become less bad may seem a double misfortune.

But how, it might be asked, could it be better to suffer a worse death if the badness of death is simply a function of the deprivation which the victim suffers? How, in other words, can it be better to be deprived of more? The proper response to these questions is, it seems, to note that there are two ways of assessing the deprivation involved in death. My claims in Section II about the comparative badness of death in infancy and death in later life were relativized to persons-at-times. But the claims of this section, which seem to clash with those earlier claims, are of a different sort: they are claims about how death affects a life as a whole. The claim that death at thirty-five is worse than death in early infancy is a claim about the relative badness of death to persons at the time of death; while the claim that it is worse for a person to die in early infancy

than at age thirty-five is a claim about the effect of death on the value of a person's life as a whole. Once we distinguish these two forms of evaluation, we can see that the two claims are in fact compatible and that there is no real paradox in the idea that it can be better for a person to suffer a worse death.