

# The Time-Relative Account of Interests

Jeff McMahan

*Guest essays represent only the views of the author(s).*

## Table of Contents

1. Introduction: Utilitarianism and Interests
2. Egoistic Concern
3. The Misfortune of Death
4. The Divergent Lives Problem
5. The No-Difference View
6. Interests and Reasons
7. An Asymmetry Between Benefits and Harms
8. Resources and Further Reading

## Introduction: Utilitarianism and Interests

“Utilitarianism” does not refer to a single moral theory. The label can be appropriately applied to innumerable different theories. Act utilitarianism and rule utilitarianism are, for example, quite different theories. One common feature of most versions of utilitarianism is that they are [consequentialist](#) theories—theories that hold that only the consequences of acts matter morally—which assume that [well-being](#) is the sole good consequence and that ill-being is the sole bad consequence, or evil. This understanding can be

subdivided into many variants. Most require some form of maximizing of well-being (total, average, minimum level, etc.). One important way in which the subdivisions are distinguished is by whether they evaluate the outcomes of acts in impersonal or in “individual-affecting” terms.<sup>1</sup>

Impersonal evaluation is concerned with the extent to which an outcome is good or bad in itself and thus with the contribution that the consequences of the act make to the value of the world. Individual-affecting evaluation, by contrast, is concerned with the effects that acts have on individuals.

Individual-affecting modes of evaluation can be further divided between “narrow” versions, which are concerned with effects in individuals *for better or worse*, and “wide” versions, which are concerned with effects on individuals that are *good or bad* for those individuals but not necessarily better or worse. These two versions come apart in the case of [acts that cause individuals to exist](#), as causing an individual to exist can be good or bad for that individual, but not better or worse. Causing an individual to exist cannot be better or worse for that individual, for the claim that it could implies that the alternative would have been worse or better for that individual—but there are no nonexistent individuals for whom never existing is worse or better than existing.

Many utilitarians have favored a [narrow individual-affecting version](#) of the theory. These utilitarians are, in the words of Jan Narveson, “in favor of making people happy, but neutral about making happy people.”<sup>2</sup> Others believe, rightly in my view, that there are moral reasons of all three sorts—impersonal, narrow individual-affecting, and wide individual-affecting—but that in some instances the narrow individual-affecting reason to produce a certain amount of well-being is stronger than the impersonal or wide individual-affecting reason to produce an equivalent amount of well-being. That is, they believe that it sometimes matters morally whether the infliction of a harm or the failure to confer a benefit is *worse for* some individual.

There are, however, different ways in which the effect of an act can be worse for an individual. These correspond to different ways of understanding the concept of an *interest*—in the sense of what is “in one’s interests” rather than what one is “interested in.” The concept of an interest has long been central to utilitarian thought. For example, Peter Singer, probably the most prominent contemporary utilitarian, has for decades articulated both utilitarianism’s radical impartiality and its egalitarian dimension by invoking this concept. His “principle of [equal consideration of interests](#)” is that “we give equal weight in our moral deliberations to the like interests of all those affected by our actions.”<sup>3</sup>

In what follows, I will be defending one way of understanding the notion of an interest. As we will see, I believe that it is through the acceptance of this understanding of an interest that moral theorists, including utilitarians, can best articulate and defend certain highly important moral truths.

## Egoistic Concern

The traditional view of prudence or self-interest is that one has reason to care about one’s future well-being in a special, egoistic way because the subject of that future well-being will be *oneself*. In other words, it is *identity* that makes egoistic concern about the future rational. Identity is, to use the term favored by Derek Parfit, *what matters*.

In what I believe to be one of the great philosophical insights of the 20th century, Parfit rejected this traditional view. His argument is based on a hypothetical example:

**Division:** Person  $P_1$ , whose body is fatally injured, is one of a set of identical triplets. The cerebral hemispheres of his two brothers,  $P_2$  and  $P_3$ , have been destroyed, though their brain stems and bodies remain alive and functional. A surgeon separates  $P_1$ ’s two cerebral hemispheres and transplants one into  $P_2$ ’s head and the other into  $P_3$ ’s head. Each

hemisphere remains conscious throughout the procedure. After a short period of consciousness without sight, sound, or sensation, each of these two centers of consciousness finds itself in a new body, though one very similar to  $P_1$ 's.<sup>4</sup>

It is universally recognized that a person can survive the loss of an entire cerebral hemisphere. If  $P_1$ 's right hemisphere were to be destroyed and his left hemisphere transplanted into  $P_2$ 's body without losing consciousness, it seems that  $P_1$  would survive in  $P_2$ 's body. And if  $P_1$ 's left hemisphere were to be destroyed and his right hemisphere transplanted into  $P_3$ 's body,  $P_1$  would survive in  $P_3$ 's body. But it does not seem that  $P_1$  can survive in Division in *both*  $P_2$ 's and  $P_3$ 's bodies; for  $P_2$  and  $P_3$ , even after the transplants, are different people—not one person with two bodies and two minds. Parfit concludes that  $P_1$  ceases to exist, but that most of what matters to him in continuing to exist is preserved in his relation to  $P_2$  and in his relation to  $P_3$ . What happens to  $P_1$  in Division is *relevantly like* double survival.

Parfit says that what this example shows is that *identity is not what matters*—or, as I will say, that identity is not the basis of rational egoistic concern about the future. Rather, what makes egoistic concern about different times rational are the relations that are the constituents of personal identity over time.

In his influential early work, Parfit argued that these relations are psychological connections and psychological continuity. Examples of psychological connections include an experience and a later memory of it, the formation of a desire and the later satisfaction or frustration of that desire, a belief at an earlier time and that same belief at a later time, and the persistence of a character trait or disposition over time. There is psychological connectedness between a person at an earlier time and a person at a later time when such psychological connections hold between them. Psychological continuity consists in overlapping chains of psychological connections over

time. There is, for example, psychological continuity between a person at times  $t_1$  and  $t_3$  when the person can, at  $t_2$ , remember an experience at  $t_1$  and, at  $t_3$ , can remember experiences at  $t_2$ , even though he cannot then remember the experience at  $t_1$ .

My view differs from Parfit's in two ways. First, Parfit thought that, provided that a later mental state is sufficiently qualitatively similar to an earlier mental state, the two states constitute a psychological connection no matter what sort of causal connection there might be between them—or, indeed, even if they are not causally connected at all.<sup>5</sup> On this view, a mental state in one brain and a later, very similar mental state in another brain constitute a psychological connection. I, however, use the term in such a way that two mental states constitute a psychological connection only when they are both states of the same functional brain matter. I think, in other words, that the relations that are the essential constituents of personal identity, and are also the basis of egoistic concern about the future, are physical, functional, and organizational continuity of the areas of the brain in which consciousness is generated. Thus, on Parfit's view, if my body were painlessly destroyed and a perfect replica of me were simultaneously created, this would be just as good for me as it would be if my body were not destroyed and no replica were created. On my view, by contrast, the prospect of my immediate replacement by a replica would eliminate any basis for my being concerned about the future in an egoistic way. I might care about what would happen to my replica for various reasons, but none of them would be relevantly like the reasons I have now to care about my own future.

The second difference is that I think there is a further relation that can ground at least a minimal degree of egoistic concern. Suppose that while my brain remained conscious, all the elements of my psychology—my memories, beliefs, desires, dispositions, and so on—would be erased, leaving something like a Lockean *tabula rasa*.<sup>6</sup> If I knew before this process began that my body would be subjected to torture immediately after the process had been

completed, it seems that it would be rational for me to fear the torture, even though there would be no psychological connections and no psychological continuity between my conscious mind before the process and the conscious mind that would remain after the process had been completed. The only relation that would hold between the mind before and the mind after the process is what I call “continuity of consciousness” (which, more precisely understood, is really continuity of the *capacity* for consciousness, in that it can include periods in which the brain is unconscious but retains the capacity for consciousness). Although this relation alone seems sufficient for egoistic concern about extreme suffering in the immediate future, it seems to ground only much weaker, if any, egoistic concern about future benefits and, in particular, about benefits or harms that would occur in the more distant future.

We can refer to the three types of relation that ground rational egoistic concern —psychological connections, psychological continuity, and continuity of consciousness—as the “prudential unity relations.” In actual cases, these relations hold only within the life of a single person, though in Division they hold between  $P_1$  before the transplants and both  $P_2$  and  $P_3$  after the transplants. Because these relations can diverge from identity only in hypothetical cases such as Division, one may wonder why, in actual cases, there is any need to distinguish between identity and the relations that are the constituents of identity. Why not just accept that, in all actual cases, identity is what matters?

Parfit supplied the answer, which is that while identity is all-or-nothing, the prudential unity relations are matters of degree. One can be strongly related by the prudential unity relations to oneself at some point in the future or only very weakly so related. Thus, whereas if identity were the basis of egoistic concern, it would be rational to be equally egoistically concerned about all parts of one’s future, if the prudential unity relations are the basis of egoistic concern, it can be rational to be much more egoistically concerned about

oneself at times when one would be strongly related to oneself by these relations than at times when one would be only weakly so related to oneself.

This last claim forms the basis of an account of the nature of interests, according to which the strength of a person's interest at time  $t$  in having some benefit in the future, or avoiding some harm, is a function of two variables: (1) the magnitude of the benefit or harm at the time when it would occur and (2) the strength of the prudential unity relations between the person at  $t$  and herself at the time the benefit or harm would occur.

The prudential unity relations normally weaken with the passage of time. That is, a person at  $t_1$  is normally less strongly psychologically connected and continuous with herself as she will be at a much later time,  $t_3$ , than with herself as she will be at an earlier time,  $t_2$ . Thus, her interest at  $t_1$  in having a benefit at  $t_2$  is normally stronger than her interest at  $t_1$  in having the same benefit at  $t_3$ . We can refer to this understanding of interests as the "Time-Relative Account of Interests," or TRAI.

## The Misfortune of Death

This general account of the nature of interests offers what I think are the best explanations of various common-sense beliefs, such as beliefs about the misfortune of death at different stages of life. It provides, for example, the basis of what I have elsewhere referred to as the "Time-Relative Interest Account" of the misfortune of death (henceforth the TRIAD), which is one way of understanding what is now the most widely accepted view of the misfortune of death: the "Deprivation Account."<sup>7</sup> According to this view, death is a misfortune for a person because it deprives that person of good life the person would otherwise have had. On the most obvious way of understanding the Deprivation Account, the *extent* to which death is a misfortune is proportional to the amount of good life of which a person is deprived by death. I have called this the "Life Comparative Account" (henceforth the LCA) because it compares

the life a person has that ends in a certain death with the life the person would have had in the absence of that death and then measures the misfortune of death by the extent to which the shorter life is less good than the longer life would have been.<sup>8</sup>

The problem with the LCA, however, is that it implies that the worst time for an individual to die is immediately after that individual begins to exist. Thus, assuming—as virtually everyone does—that we begin to exist at some point before birth, this account implies that the death of a fetus is a substantially greater misfortune for the fetus than the death of a child, adolescent, or young adult is for that individual. But that implication is deeply implausible.

The TRIAD avoids this implication by claiming that the misfortune of death for an individual is a function not just of the amount of good life that the individual loses through death but also of the strength of the prudential unity relations that would have held between the individual at the time of death and that same individual at later times if the death had not occurred. According to the TRIAD, death immediately after an individual begins to exist (whether that is at conception or at some other time during fetal gestation) is hardly a misfortune at all for that individual, since the individual at the time of death would have been connected to itself later, when it would experience benefits and harms, only by continuity of consciousness, if that. The life that an individual loses by dying shortly after beginning to exist is, as Parfit might have said, *relevantly like* the life of a different individual. Hence when an individual dies immediately or shortly after beginning to exist, that is *relevantly like* that individual's never coming into existence at all. On the TRIAD, death gradually becomes an increasingly greater misfortune for an individual as that individual develops and matures psychologically—a process that normally accelerates dramatically shortly after birth—even though as this development proceeds, the amount of remaining good life is gradually decreasing.

The TRIAD, which is implied by the TRAI, supports a permissive view of the morality of abortion, as it implies that death is not a significant misfortune for a fetus. If, as many people believe, we begin to exist before the fetal brain develops the capacity for consciousness, the death of a fetus before this point may not be a misfortune at all for that fetus, as no prudential unity relations would hold at all between the fetus and itself in the future. And if, as others believe, we begin to exist when, or even after, the capacity for consciousness arises, there would be psychological connections and psychological continuity of only negligible strength between the conscious fetus and itself in the future when it would be capable of having a substantial level of well-being. Fetuses, therefore, have either no interest in continuing to live or, later in the course of gestation, only a very weak interest in continuing to live. On this account, then, apart from any side effects involving the interests of others, the killing of a fetus frustrates at most only a very weak interest, and thus inflicts only a comparatively insignificant harm.

There are, however, instances in which painlessly killing an individual seems seriously wrong even though the individual has no interest in continuing to live or only a very weak interest in continuing to live—for example, the killing of a person who cannot live longer than another day and will do nothing but lie in bed until he dies. Yet the wrongness of killing such an individual is not explained by the frustration of his interests, but by the failure to respect him as an autonomous person whose nature and capacities ground his moral status as an individual with rights, or an individual whose treatment is governed by moral constraints. And this form of explanation of the wrongness of killing seems inapplicable to fetuses that either lack the capacity for consciousness or have only the capacity for a form of consciousness that is even more rudimentary than that of many nonhuman animals. The challenge for those who believe that fetuses have a moral status that makes killing them seriously wrong, even though it barely harms them, is to explain why, for example, adult mammals [lack the same status](#) (assuming, as seems warranted in most cases,

that these people think the killing of a fetus is more seriously wrong than the killing of a rat).

There is, however, a problem here. If the absence, or paucity, of prudential unity relations between a fetus and itself in the future means, as the TRAI implies, that it has at most only very weak interests in what might happen to it later in life, it seems that a fetus can also have only a very weak interest in not being caused a painless prenatal injury, even if that injury would make its future life significantly less good. And that may suggest that there is no reason, grounded in the fetus's interests, not to injure or damage it.

The obvious response to this objection is to note that, whereas abortion frustrates at most only a weak present interest in continuing to live and also prevents the fetus from ever having any other interests, a non-lethal prenatal injury may frustrate many strong interests that the person who develops from the fetus will have over the course of an entire life. If this is right, the TRAI can both help to explain the permissibility of abortion in many or most cases and also explain the general impermissibility of inflicting prenatal injury.

## **The Divergent Lives Problem**

I once thought that this is the correct way to explain the presumptive wrongness of inflicting prenatal injury. While I still believe that this is the correct explanation in some cases, I now think it fails in many other cases. Suppose that a woman is advised, early in pregnancy, to take folic acid supplements to reduce the risk of her child's developing spina bifida (a birth defect in which part of the spine protrudes through the baby's back). She refuses and also indulges in daily saunas, which she knows increase the risk of spina bifida. Later she gives birth to a child with a serious case of spina bifida. Suppose that, if she had taken supplements and not taken saunas, the same child would have been born but without spina bifida.

The infant, because of her condition, requires a variety of special treatments beginning immediately after birth and continuing for the rest of her life. These

treatments, together with the limits imposed by her condition on the activities in which she can engage, cause the events in her life, her experiences, and indeed her own character and concerns to diverge radically from the way they would have been had she not developed spina bifida. In her actual life with spina bifida, she will meet different people from those she would have met in the absence of the condition. She will engage in different activities, pursue a different vocation, have different values and aims, marry a different person, and have different children. And these particular attachments, ambitions, and values will lead to the development of different *interests*. In her actual life, her distinctive interests (in contrast to her generic interests, such as the interest in avoiding suffering) will be concerned with her actual aims and ambitions, and with her relations with her actual partner, children, and friends—most or all of which, or whom, would have been different if she had not had spina bifida. Thus, most of the *objects* of her actual interests would not have been part of her life if she had not had spina bifida. Her *actual interests* are therefore ones she would not have had if she had not developed spina bifida. Because of this, the interests she has in her life with spina bifida are almost certain to be better satisfied, on balance, in her actual life than those *same interests* would have been in a life in which she never developed spina bifida. Little of what she actually—and rationally—cares about would have been present in that alternative life, and she herself would have been quite different in ways that she would find alien. If, as an adult, she could know what she would have been like at the same age in that alternative life, she might reasonably, though figuratively, think “That would not have been *me!*” There may be no time in her life after she has developed substantial interests when it would be rational for her to regret having been caused to have spina bifida.

We can refer to the fact that a person can rationally prefer her actual, less good life to a better life she could have had as the “Divergent Lives Problem.” This problem is in certain respects an intrapersonal analogue of Parfit’s well-known [Non-Identity Problem](#). Suppose that a woman is advised by her physician to postpone conception for several months and take folic acid

supplements during that period to reduce the risk of having a child with spina bifida. She ignores that advice and the fetus she conceives develops spina bifida because she in fact had a serious folate deficiency. This child has a less good life than a different child would have had if the woman had taken supplements for three months and then conceived that different child, who would not have had spina bifida. But as long as the actual child with spina bifida has a life that is worth living, she has no egoistic reason to regret that she exists rather than that other, better-off child. Similarly, my claim about the original case, in which the child with spina bifida could herself have had a life without it, is that that alternative would have been *relevantly like* the life of a different person. And that is why the child with spina bifida may have no egoistic reason to regret that she has the life she has rather than an alternative, impartially better life without spina bifida.

The Divergent Lives Problem arises to varying degrees. It arises, for example, in a relatively weak form when someone inflicts a harm on an adult that changes the course of her life in ways that cause her to meet different people, become interested in things that she would not otherwise have come to care about, and so on. This phenomenon is pervasive and is of substantially less moral significance than it is when it arises in its strongest form, as in the example involving spina bifida. This is because the adult, unlike the fetus that develops spina bifida, has strong interests at the time she suffers the harm and these interests are frustrated by that harm. The spina bifida example is an instance of the Divergent Lives Problem in its strongest form because the effects that adversely affect the child's life begin very early and are extensive, though the *significantly bad* effects begin to occur only after the victim has (i) begun to develop a distinctive character, (ii) established attachments to particular people, and (iii) come to value certain activities rather than others.<sup>9</sup>

Because of the Divergent Lives Problem, the mother's having caused her child to have spina bifida was not against the child's actual interests, on balance. Hence it is not true that the mother's action was objectionable because it

would frustrate many of the child's most significant later interests. The mother's action seems clearly objectionable, but the explanation is not that the action was against her child's interests—at least if interests are understood as the TRAI understands them.

There is, however, another, more familiar way of understanding interests. I mentioned earlier that the Life Comparative Account of the misfortune of death offers a different construal of the Deprivation Account from that given by the TRIAD. And just as the TRIAD is implied by the more general TRAI, so the LCA is implied by what can be called the “Life Comparative Account of Interests” (LCAI). According to this account, an individual has an interest in whatever would provide the best life possible, impartially considered, for that individual, irrespective of the strength of the prudential unity relations within the individual's life at different times. As soon as an individual begins to exist, that individual has an interest in having the best life possible—an interest that subsumes all more particular interests in having particular benefits and avoiding particular harms. And that interest is stronger than any of the interests the individual will later have. This is because it is just after an individual begins to exist that there is the most good in prospect. The amount of future well-being that is possible at this point will never be possible again; for, as the individual's life continues, there is less and less time remaining in which the life might be either better or less good. Hence the individual's interest in having the best life possible necessarily becomes weaker and weaker as time passes.

On the LCAI, the mother's action in causing her child to have spina bifida frustrates the interest the child has in having the best life possible. Thus, the infliction of prenatal injury may, depending on the severity of the injury, frustrate a stronger interest than those frustrated by many or most acts that inflict greater injuries later in life. This is true even in cases in which the Divergent Lives Problem arises in a strong form, so that the individual's later interests are better satisfied in the life with the prenatal injury than those same interests would have been in a life without it. For, at the time that the

prenatal injury is inflicted, the all-subsuming interest the fetus has in having the best possible life is immensely strong. And the fetus has no distinctive *future* interests, as what interests it will develop in the future (other than generic interests such as the interest in avoiding suffering) depends on whether or not the prenatal injury is inflicted. And, to the extent that the fetus's possible interests matter morally, the individual is likely to have more and stronger interests that would be better satisfied overall in a life without the injury than in a life with it.

But, while the LCAI thus offers an explanation of the wrongness of inflicting prenatal injury, it may overestimate the seriousness of the wrong. More importantly, it implies that the killing of a fetus that would otherwise have been identical with a later adult person inflicts a greater harm on that individual than the killing of that same individual would inflict at any later time. While this implication does not entail that killing an individual immediately after it comes into existence is more seriously wrong than killing the same individual at a later time, it does strongly suggest that such a killing—which frustrates so strong an interest—must be seriously wrong. Yet even those who believe that early abortions are seriously wrong tend to base their view, not on the claim that the fetus suffers a terrible harm in being killed, but on claims about its high moral status. This is shown by the fact that most opponents of abortion are far less concerned about the fact that more than a quarter of pregnancies end in spontaneous abortion during the first 20 weeks. They also tend, when they learn of the spontaneous abortion of an early fetus, not to react with the same emotions they experience when they learn of the death of, for example, a 12-year-old child.

If the TRAI is the more plausible account, it is more difficult to explain the wrongness of inflicting prenatal injury, at least in cases in which the Divergent Lives Problem arises in a strong form. Because the fetus at the time of the injury would not be connected to any significant degree to itself in later life by the prudential unity relations, the infliction of prenatal injury is relevantly like causing a less well-off person to exist rather than causing or allowing a

different, better-off person to come into existence. This parallels my earlier claim that having an abortion, and particularly an earlier abortion, is relevantly like preventing a person from coming into existence. If it is right that the wrongness of prenatal injury must, at least in many instances, be explained by much the same considerations that explain the wrongness of causing a less well-off rather than a better-off person to exist, this has important implications for [population ethics](#). For even when we come to understand the Divergent Lives Problem, we continue to believe that the infliction of prenatal injury is seriously wrong, and more seriously wrong the more severe the injury is (even though the more severe the injury is, the more extensive the Divergent Lives Problem is likely to be). If this belief is correct, and if inflicting a prenatal injury is relevantly like causing a less well-off person to exist, then to cause a less well-off person to exist rather than a different, better-off person must be more seriously wrong than common-sense morality has recognized.

## The No-Difference View

One might argue that there is in fact a morally significant difference between inflicting prenatal injury and causing a less well-off rather than a better-off person to exist. The difference is that, whereas the infliction of prenatal injury is *worse for* the injured individual to the extent that it causes that individual to have a less good life, causing a less well-off rather than a better-off person to exist is not worse either for the person who never exists or for the less well-off person who is caused to exist with a life worth living. It is not worse for the former because there cannot be anyone for whom never existing is worse than existing. And it is not worse for the latter because the alternative for that individual was never to exist. And the claim that existing is worse for a person than never existing implies that never existing would have been better for that person. But, again, there cannot be anyone for whom never existing is better than existing. Thus, if the reason not to cause a bad effect is stronger when that effect would be worse for someone than when it would not be worse for

anyone, then the reason not to inflict prenatal injury is stronger than the reason not to cause a less well-off rather than a better-off person to exist, assuming that the reduction in well-being would be the same in each case.

Derek Parfit rejects the claim that the reason not to cause a bad effect is stronger when that effect would be worse for someone. According to what he calls the “No-Difference View,” it makes no moral difference whether a bad effect is worse for someone or whether a good effect is better for someone. On this view, there is no moral difference between the following two acts:

(1) To secure some minor benefit for herself, a pregnant woman,  $W_1$ , knowingly inflicts an injury on her fetus that will lead to a strong form of the Divergent Lives Problem. Because of her act, the person who develops from this fetus will have a lifetime level of well-being of 1000 when he would otherwise have had a level of 1300.

(2) Another woman,  $W_2$ , has two frozen embryos but can have only one implanted. They are genetically identical except that one,  $E_1$ , has a genetic disorder that will enable the person who develops from it to have a lifetime well-being level of only 1000. The other,  $E_2$ , would produce a person whose lifetime well-being would be 1300. Unless  $W_2$  chooses otherwise,  $E_2$  will be implanted. To secure a minor benefit for herself,  $W_2$  has  $E_1$  implanted instead.

According to the TRAI, there is indeed no significant moral difference between these acts, assuming that the Divergent Lives Problem arises in its strongest form as a result of  $W_1$ 's act.

Yet, even though the No-Difference View aligns with the TRAI in its implications about these examples, there are other comparisons about which its implications are difficult to believe (which is not to suggest that its implications for the preceding comparison are intuitive). It implies, for

example, that there is also no moral difference between the following two options. Suppose that one can do either but not both of two acts:

(3) One can save the life of a 60-year-old stranger who will then live to 80. This would, however, require a significant personal sacrifice and so is supererogatory.

(4) Unless one intervenes, a person will come into existence who will die at age 60. One can, without harming anyone or violating anyone's rights, prevent this person from coming into existence and cause a different person to exist instead whose quality of life would be the same but who would die at 80. This would require the same sacrifice as saving the 60-year-old would require and so is also supererogatory.

Although both acts are supererogatory, so that failing to save the existing person would not wrong him or violate his rights, most people would, I believe, think that the moral reason to save the 60-year-old is stronger than the reason to cause the longer-lived person to exist rather than allow the shorter-lived person to exist instead. Yet both acts would have the good effect of enabling someone to live 20 years *more*—in the first instance, more than *the person himself* would otherwise live and, in the second instance, more than *someone else* would otherwise live. The reason, I suggest, that it matters more to enable the existing person to live more years than he would otherwise live is that this would be *better for* him, so that failing to save him would be *worse for* him. By contrast, if one fails to cause the longer-lived person to exist and allows the shorter-lived person to exist instead, this will not be *worse*, or bad, for either. If anything, it would be *good* for the shorter-lived person, who would receive the “existential benefit” of a life worth living.

Unless we are willing to accept that the moral reason to save the 60-year-old is no stronger than the reason to cause the longer-lived person to exist, we must reject the No-Difference View. And this is equivalent to accepting that, at least in many cases, it makes a moral difference both whether a bad effect is worse for the victim and whether a good effect is better for the beneficiary.

Yet this is inconsistent with the implication of the TRAI that making a person's life as a whole less good by amount  $x$  through the infliction of prenatal injury is, at least when the Divergent Lives Problem arises in a strong form, relevantly like causing a person to exist whose life will be less good by amount  $x$  than the life of another person one could have caused to exist instead. For the prenatal injury is worse for the injured individual, while causing the shorter-lived person to exist is not worse for anyone. If it makes a moral difference that the infliction of prenatal injury is worse for the injured individual, the infliction of the prenatal injury should be more seriously wrong—and that, I concede, is what most people believe.

The explanation of this inconsistency is that, according to the TRAI, there are two morally distinct (though usually overlapping) ways in which a bad effect can be worse for someone. If one fails to save the 60-year-old stranger, that is worse for him in that it frustrates (or allows the frustration of) strong interests he has in having the benefits of an additional 20 years of good life—years in which he would have been strongly connected to himself at 60 by the prudential unity relations. But if, by contrast, one inflicts prenatal injury on a fetus and this leads to a strong form of the Divergent Lives Problem, this will be very unlikely to frustrate strong interests that the individual will have. Indeed, as I have argued, it is very likely that the act will *better* satisfy the interests the individual *will* have. In this sense, which we can call the “interest-based sense,” the prenatal injury was *not* worse for her. Thus, if she could know, at any time in her life, what her life at that point would have been like without the injury, she rationally could, and probably would, be glad to have her actual life with the injury, with her actual partner, children, and friends, her actual character and values, and her actual ambitions and achievements.

The prenatal injury was, of course, worse for her in what we can call the “life-comparative sense,” in that it caused her to have a less good life as a whole, impartially considered. But recall that the fetus at the time of the injury would have been equally unconnected by the prudential unity relations to itself in the

less good life and to itself in a better life. And, as I have just noted, if in her actual life with the injury, the individual could know what both she and her circumstances would have been like in the counterfactual life without the injury, she would be likely to regard herself in that life as relevantly like a different person, and to regard the life itself as relevantly like the life of a different person. Thus, that the prenatal injury was worse for her in the life-comparative sense lacks the moral significance of the fact that the failure to save the 60-year-old would be worse for him in the interest-based sense.

In summary, the comparison between failing to save an existing person and allowing a shorter-lived person to exist rather than causing a different, longer-lived person to exist shows that it can matter morally whether a bad effect is worse for someone. But a further comparison between failing to save an existing adult and the infliction of prenatal injury that results in a strong form of the Divergent Lives Problem compels us to refine that initial conclusion by accepting that what really matters is whether a bad effect is worse for someone in the interest-based sense, not whether it is worse for someone in the life-comparative sense. In short, the explanation of why it matters that a bad effect is worse for someone is that it matters whether interests are frustrated. The reason not to cause a bad effect that frustrates an interest is stronger than the reason not to cause an equivalent bad effect when doing so does not frustrate any interest.

## Interests and Reasons

This last claim presupposes that the TRAI is the correct account of interests. If the LCAI were true, causing an individual to have a less good life as a whole would always frustrate a strong interest, even in cases involving the strongest form of the Divergent Lives Problem. In that case, there would be no distinction between an act that is worse for an individual in the interest-based sense and one that is worse for a person in the life-comparative sense. I have, however, offered arguments to show that the TRAI and not the LCAI is the correct account of interests.

In Section 1, following Parfit, I distinguished three types of moral reason: impersonal, narrow individual-affecting, and wide individual-affecting. In accepting the No-Difference View, Parfit accepted that no one of these types of reason is in general stronger than either of the others. But, as I noted earlier, most of us believe that, in some cases at least, narrow individual-affecting reasons are stronger than either of the other two types of reason. The reason to save the 60-year-old is, for example, a narrow individual-affecting reason, whereas the reason to cause the longer-lived person to exist rather than allow the shorter-lived person to exist is either impersonal or wide individual-affecting. If one causes the longer-lived person to exist, that will be *good for* that person by more—20 years of good life more—than causing the shorter-lived person to exist would be good for that person. But saving the 60-year-old would be *better for* that person—also by 20 years of good life—and that matters more.

What I have argued here is that the explanation of why narrow individual-affecting reasons are stronger in some cases is that the failure to act on such a reason will be worse for some individual, and it will be worse for that individual because it frustrates one or more of that individual's interests. When the failure of an individual to receive some benefit is worse for that individual because it frustrates one of that individual's interests, that matters more morally than the non-occurrence of an equivalent benefit when that would be bad only impersonally, or the non-occurrence of a benefit that would have been good but not better for an individual (which is the case when a person is benefited by being caused to exist).

Strictly speaking, the reason not to inflict prenatal injury when the Divergent Lives Problem would be present in its strongest form is a narrow individual-affecting reason. But in this case the narrow individual-affecting reason lacks its usual strength. In substance, it is relevantly like an impersonal or wide individual-affecting reason. This is because the injury would be worse for the individual only in the life-comparative sense; and it is only when an effect is worse for an individual in the interest-based sense that the reason not to

produce that effect is stronger than a corresponding impersonal or wide individual-affecting reason.

Reasons of these latter two types necessarily take no account of interests. It is only narrow individual-affecting reasons that are concerned with interests. When interests are understood in the way the TRAI understands them, they explain why an act's being worse or better for an individual has special moral significance, and thus explain why narrow individual-affecting reasons are, contrary to the No-Difference View, normally stronger than corresponding impersonal or wide individual-affecting reasons. It is thus the TRAI that explains the common-sense belief that narrow individual-affecting reasons are normally stronger than other types of reason, even when the same amounts of well-being are at stake.

## **An Asymmetry Between Benefits and Harms**

I will conclude with one highly important qualification. This is that the moral significance of interests seems greater in the case of interests in having benefits than in the case of interests in avoiding intrinsic harms, such as suffering. In the choice between saving the 60-year-old and causing the longer-lived person to exist, what is at stake are benefits: the benefit to the 60-year-old of having 20 more years of good life and the greater existential benefit to the longer-lived person rather than the substantially lesser existential benefit to the shorter-lived person. In this choice, the reason to confer a benefit when that would be better for the beneficiary is stronger than the reason to provide an equivalent benefit that would not be better for the beneficiary.

But now consider a choice between (1) preventing an existing person from experiencing suffering of great intensity and duration and (2) causing a person to exist whose life would not contain such suffering rather than allowing a different person to exist whose life would be much the same except that it would contain suffering of the same intensity and duration that one could

prevent in the life of the existing person. The reason to prevent the suffering of the existing person is a narrow individual-affecting reason while the reason to cause the person to exist whose life would not contain the suffering rather than the person whose life would contain it is either impersonal or wide individual-affecting. But in this choice, it does not seem so obvious that the narrow individual-affecting reason is stronger—or, perhaps, stronger to the same extent. It may be somewhat stronger, but the difference in strength between the narrow individual-affecting reason and the other type of reason does not seem very great.

It is not that the interest in avoiding an intrinsic harm of a certain magnitude is stronger than the interest in having a benefit of the same magnitude; rather, it seems to be that it matters less, perhaps much less, whether an intrinsic harm is worse for an individual than whether the absence of a benefit is worse for an individual.

## About the Author



**Jeff McMahan** is Sekyra and White's Professor of Moral Philosophy at the University of Oxford and a Fellow of Corpus Christi College. He is the author of *The Ethics of Killing: Problems at the Margins of Life* (Oxford University Press, 2002) and *Killing in War* (OUP, 2009). He is currently writing a book, *The Ethics of Creating, Preserving, and Ending Lives*, which will also be published by OUP.

## How to Cite This Page

McMahan, J. (2023). The Time-Relative Account of Interests. In R.Y. Chappell, D. Meissner, and W. MacAskill (eds.), *An Introduction to Utilitarianism*, <<https://www.utilitarianism.net/guest-essays/time-relative-account-of-interests>>, accessed 8/12/2025.

# Resources and Further Reading

- Robert Merrihew Adams, “Existence, Self-Interest, and the Problem of Evil”
- Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (New York: Oxford University Press, 2002), ch. 2.
- \_\_\_\_\_ “Paradoxes of Abortion and Prenatal Injury,” *Ethics* 116, 4 (2006): 625–55.
- \_\_\_\_\_ “The Lucretian Argument,” in R. Feldman, K. McDaniel, J.R. Raibley, and M.J. Zimmerman, eds., *The Good, the Right, Life and Death: Essays in Honor of Fred Feldman* (Aldershot, UK: Ashgate Publishing, 2006): 213–26.
- \_\_\_\_\_ “Creating People and Saving People,” in Gustaf Arrhenius, Tim Campbell, and Krister Bykvist, eds., *The Oxford Handbook of Population Ethics* (Oxford: Oxford University Press, 2022): 51–45.
- Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1987 printing)
- David Velleman, “Persons in Prospect,” *Philosophy & Public Affairs* 36 (2008): 221–88.

- 
1. I substitute this term for the one that Derek Parfit coined – “person-affecting” – in order to emphasize that this mode of evaluation applies to effects on animals and human non-persons as well as to persons. See Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1987 printing), p. 370. ↩
  2. Jan Narveson (1973). “Moral Problems of Population,” *The Monist*, 57: 62–86, p. 80. ↩

3. Peter Singer, *Practical Ethics*, Third Edition (New York: Cambridge University Press, 2011), p. 20. ↩
4. Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1987 printing), pp. 253–66. ↩
5. The idea that two causally unrelated mental states can form a psychological connection is not, to the best of my knowledge, stated anywhere in Parfit’s published work. But I argued later in his life that his reductionist view of what matters should not require causation and he confirmed in personal correspondence that he had changed his view to accept that claim. ↩
6. Bernard Williams, “The Self and the Future,” in *Problems of the Self* (Cambridge: Cambridge University Press, 1973). ↩
7. The classic statement of the Deprivation Account is in Thomas Nagel, “Death,” in *Mortal Questions* (Cambridge: Cambridge University Press, 1979). The TRIAD is defended in Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (New York: Oxford University Press, 2002), ch. 2. ↩
8. *The Ethics of Killing*, p. 105. ↩
9. Compare Parfit, *Reasons and Persons*, p. 523, footnote 16, in which he responds to early articulations of the Divergent Lives Problem by suggesting that his case of the Medical Programmes could involve only delayed harms. I explore the Divergent Lives Problem in much greater detail in draft articles on, among other things, prenatal injury, wrongful life, and gene editing, which I hope will appear in print soon. ↩