

DEREK PARFIT

## Future Generations: Further Problems

There are three kinds of choice. *Different Number Choices* affect both the number and the identities of future people. *Same Number Choices* affect the identities of future people, but do not affect their number. *Same People Choices* affect neither.

Different Number Choices raise well-known problems. Less well-known are the problems raised by Same Number Choices. Both are discussed above by Gregory Kavka.<sup>1</sup> I shall continue this discussion.<sup>2</sup>

### I. THE FUTURE INDIVIDUALS PARADOX

Consider first a Same People Choice:

*The Nuclear Technician.* Some technician lazily chooses not to check some tank in which nuclear wastes are buried. As a result there is a catastrophe two centuries later. Leaked radiation kills and injures thousands of people.

We can plausibly assume that, whether or not this technician checks this tank, the same particular people would be born during the next two centuries. On this assumption, this technician's choice is worse for the people struck by the catastrophe. If he had chosen to check

1. This issue; pp. 93–112.

2. This paper derives from a typescript, "Overpopulation: Part One," that I circulated in 1976. This is why I do not refer to those who have later reached and published similar conclusions. Throughout this paper "he" means "he or she."

the tank, these same people would have later lived, and escaped the catastrophe.

Is it morally relevant that the people whom this technician harms do not yet exist when he makes his choice? On one view, moral principles only cover those who can reciprocate—those who can benefit or harm each other. Since these future people cannot affect the nuclear technician, the harm that he causes them has, on this view, no moral significance. I shall assume that we should reject this view.<sup>3</sup> According to another view, while we ought to be concerned about the effects of our acts on future people, we are justified in being less concerned about more remote effects. We can *discount* such effects at some rate of  $n$  per cent per year. I shall assume that this view should also be rejected. What we may accept is that we have special obligations to some of our contemporaries, and that, for this reason, we ought to be more concerned about some presently existing people than about most future people. But our concern about future people should not be less simply when, and simply because, they will live later. If other things are equal, we ought to have as much concern about the predictable effects of our acts whether these will occur in 200 or 400 years. This has great importance. Nuclear wastes may be dangerous for thousands of years. And some of our acts have permanent effects.<sup>4</sup>

Consider next:

*The Risky Policy.* Suppose that, as a community, we have a choice between two energy policies. Both would be completely safe for at least two centuries, but one would have, for the further future, certain risks. If we choose the Risky Policy, the standard of living would be slightly higher over the next century. We do choose this policy. As a result there is a similar catastrophe two centuries later, which kills and injures thousands of people.

Unlike the Nuclear Technician's choice, our choice between these policies affects who will be later born. I shall expand Kavka's explanation.

3. For the reasons given by Brian Barry's "Circumstances of Justice and Future Generations," in *Obligations to Future Generations*, ed. R. I. Sikora and Brian Barry (Philadelphia: Temple University Press, 1978).

4. I argue briefly against the *Social Discount Rate* in my "Energy Policy and the Further Future," in *Energy and the Future*, ed. Douglas Maclean and Peter G. Brown, Maryland Studies in Public Philosophy, forthcoming.

Our identity in fact partly depends on when we are conceived. This is so on both of the main views about this subject. Consider some particular person, such as yourself. You are the *n*th child of your mother, and you were conceived at time *t*. According to one view, you could not have grown from a different pair of cells. If your mother had conceived her *n*th child some months earlier or later, that child would in fact have grown from a different pair of cells, and so would not have been you.

According to the other main view, you could have grown from different cells, or even had different parents. This would have happened if your actual parents had not conceived a child when they in fact conceived you, and some other couple had conceived an extra child who was sufficiently *like* you, or whose life turned out to be sufficiently like yours. On this other view, that child would have been you. But those who take this view, while believing that you *could* have grown from a different pair of cells, would admit that this would not *in fact* have happened. On both views it is in fact true that, if your mother had conceived her *n*th child at a different time, that child would not have been you, and *you* would never have existed.

Return now to the choice between our two policies. If we choose the Risky Policy, the standard of living will be slightly higher over the next century. This effect implies another. It is not true that, whichever policy we choose, the same particular people will exist two centuries later. Given the effects of two such policies on the details of our lives, it would increasingly over time be true that people married different people. More simply, even in the same marriages, the children would increasingly be conceived at different times.<sup>5</sup> As we have seen, this would in fact be enough to make them not the same children. The proportion of those later born who would owe their existence to our choice of one of the two policies would, like ripples in a pool, steadily grow. We can plausibly assume that, after two centuries, there would be no one living who would have been born whichever policy we chose. (It may help to think of this example: How many of us could truly claim, "Even if railways had never been invented, I would still have been born"?)

In my imagined case, we choose the Risky Policy. As a result, two

5. The British Miner's Strike of 1974, when television closed down an hour early, affected the timing of thousands of conceptions.

centuries later, thousands of people are killed and injured. But if we had chosen the alternative Safe Policy these particular people would never have existed. Different people would have existed in their place. Is our choice of the Risky Policy worse for anyone?

We can first ask: "Could a life be so bad—so diseased or deprived—that it would not be worth living? Could a life be even worse than this? Could it be worse than nothing, or 'worth *not* living'?" Like Kavka, I shall assume that there could be such lives. But we do not yet need this assumption. We can suppose that, whether or not lives could be worth not living, this would not be true of the lives of the people killed by the catastrophe. These people's lives will be well worth living. And we can suppose the same of those who mourn for those killed, and those whom the catastrophe disables. (Perhaps, for some of those who suffer most, the *rest* of their lives would be worth not living. But this would not be true of their lives as a whole.)

We can next ask: "If we cause someone to exist, who will have a life worth living, do we thereby benefit this person? Do we also benefit this person if some act of ours is a necessary part of the cause of his existence?" These are difficult questions. If we answer yes to both, I shall say that we believe *that causing to exist can benefit*. Since there would be many who would answer no, I shall discuss the implications of both answers.

Because we chose the Risky Policy, thousands of people are later killed or injured or bereaved. But if we had chosen the Safe Policy, these particular people would never have existed. Suppose that we do not believe that causing to exist can benefit. We can ask, "If particular people live lives that are on the whole well worth living, even though they are struck by some catastrophe, is this worse for these people than if they had never existed?" Our answer must be no. If we do believe that causing to exist can benefit, we can say more. Since the people struck by the catastrophe live lives that are well worth living, and would never have existed if we had chosen the Safe Policy, our choice of the Risky Policy is not only not worse for these people: it *benefits* them.

Let us now compare our two examples. The Nuclear Technician chooses not to check some tank. We choose the Risky Policy. Both these choices predictably cause catastrophes, which harm thousands

of people. These predictable effects both seem bad, providing at least some moral objection to these choices. In the case of the technician, the objection is obvious. His choice is worse for the people who are later harmed. But this is not true of our choice of the Risky Policy. Moreover, when we understand this case, we *know* that this is not true. We know that, even though our choice may cause such a catastrophe, it will not be worse for anyone who ever lives.

Does this make a moral difference? There are three views. It might make all the difference, or some difference, or no difference. There might be no objection to our choice, or some objection, or the objection may be just as strong.

Some claim

*Wrongs Require Victims:* Our choice cannot be wrong if we know that it will be worse for no one.

This implies that there is no objection to our choice. We may find it hard either to deny this claim, or to accept this implication. This is Kavka's *Future Individuals Paradox*.

Like Kavka, I deny that Wrongs Require Victims. If we know that we may cause such a catastrophe, I am sure that there is at least some objection to our choice. I am inclined to believe that the objection is just as strong as it would have been if, as in the Case of the Nuclear Technician, our choice would be worse for future people. If this is so, it is morally irrelevant that our choice will be worse for no one. This may have important theoretical implications. But I shall not explore these here.<sup>6</sup> The prior question is, what *is* the objection to our choice?

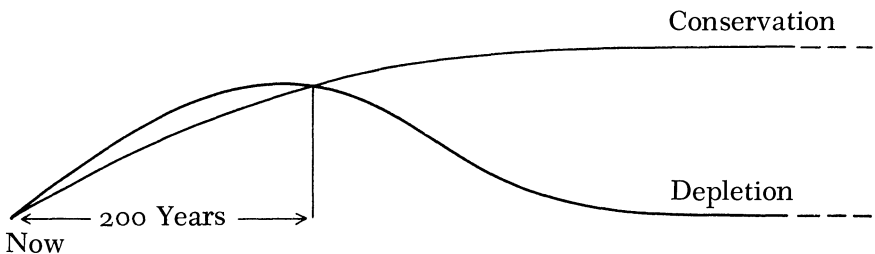
Before we discuss possible answers, it will help to introduce some more examples. We must continue to assume that some people can be worse off than others, in morally significant ways, and by more or less. But we need not assume that these comparisons could be even in principle precise. There may be only rough or partial comparability. By "worse off" we need not mean "less happy." We can be thinking, more narrowly, of the standard of living, or share of resources per person, or, more broadly, of the quality of life. And we can assume throughout that these three—the level of happiness, the

6. They are briefly sketched in the *Theoretical Footnote* to my "Energy Policy and the Further Future," op. cit.

share of resources, and the quality of life—correlate, or rise and fall together. Our arguments will thus apply whichever of these three we believe to be morally important. Since it has the broadest meaning, I shall most often use the phrase “the quality of life.” And I shall extend the ordinary use of the phrase “worth living.” If one of two groups of people would have a lower quality of life, I shall call their lives to this extent “less worth living.” I shall also assume that lives could be so bad as to be not worth living, or even to be worth not living. This last assumption is not essential to most of what follows.

Here are three more examples:

*Depletion.* Suppose that, as a community, we must choose whether to deplete or conserve certain kinds of resources. If we choose Depletion, the quality of life over the next two centuries would be slightly higher than it would have been if we had chosen Conservation, but it may later be much lower. At this much lower level people’s lives would, however, still be well worth living. The effects might be shown like this:



*The Handicapped Child.* Some woman knows that if she conceives a child now it will have some handicap. If she waits, she will later conceive a different child, who will be normal. She chooses not to wait, and knowingly conceives a handicapped child. This child’s life will be less worth living than the life of a normal child, but it will still be worth living. (These predictions would not be certain, but we can suppose that they could be reasonably assumed.)

*The 14-Year-Old Girl.* This girl chooses to have a child, and gives him a poor start in life. If she had waited several years, she would

have had a different child, to whom she would have given a better start in life.

These cases raise the same problem. We know that, if we choose Depletion, this may greatly lower the quality of life in the further future. This seems to provide at least some moral reason not to choose Depletion. But, as with the Risky Policy, we know that our choice will be worse for no one. And it seems wrong to choose to have a handicapped or disadvantaged child, when by merely waiting one could have a child free from these burdens. But these two choices will not be worse for these two children.

## II. KAVKA'S PROPOSED SOLUTIONS

What is the objection to these and similar choices? Kavka makes three proposals. His *modified Kantian imperative* forbids treating as a mere means the creation of rational beings. This principle seems plausible, especially in two of Kavka's cases. It explains what seems most wrong in having a child merely so that he can either provide a kidney for his father, or be sold into slavery.

As Kavka says, his Kantian imperative cannot be straightforwardly applied to the Future Individuals Paradox. When we choose the Risky Policy, or Depletion, we do not treat the creation of future people as a mere means. It is true that future people will in part owe their existence to our choice. This is why we know that our choice will be worse for no one. The problem is to explain why, if we know this, there is an objection to our choice. Kavka suggests that we might stretch his Kantian imperative. We might point out that, in the statement of this problem, "the fact of creation is used as a means of cancelling the prima facie moral prohibition on the acts in question that derives from their undesirable effects."<sup>7</sup> But this just assumes that these effects are undesirable. It does not explain why.

Kavka's second principle aims to explain why. He calls certain kinds of life *restricted*, and then claims that, other things being equal, it is "intrinsically undesirable from a moral point of view" that restricted lives be lived.<sup>8</sup> (The word *intrinsically* means *in itself*, and

7. p. 111.

8. p. 105.

allows that effects on other people, or other considerations, may change our overall conclusion. Thus painful surgery might be intrinsically bad but all things considered good.) Kavka's claim about restricted lives needs to be made clearer. He does not merely mean that, if someone lives a restricted life, this is worse than if this same person lived an unrestricted life. But he might mean

(1) If someone lives a restricted life, this is intrinsically worse than if this person had never existed, and someone else had existed in his place and lived an unrestricted life.

Or he might mean

(2) If someone lives a restricted life, this is intrinsically bad, worse than if this person had never existed, and no one had existed in his place.

He seems to mean (2) rather than (1). He believes that "there is something seriously wrong with people living restricted lives," and that we ought to "prevent" such lives.<sup>9</sup> He thinks it intrinsically *undesirable* that such lives be lived, not merely, as (1) claims, that this is *less desirable* than that unrestricted lives be lived.

Whether (2) is plausible depends on what counts as "restricted." Kavka calls lives restricted when they are "significantly deficient in one or more of the major respects that generally make human lives valuable and worth living."<sup>10</sup> He adds that such lives will "typically be worth living on the whole."<sup>11</sup> Is (2) plausible when applied to lives that are *well* worth living?

One example might be that of a life cut short in its prime. Consider someone whose life is well worth living, but who dies at 35. Setting aside effects on other people, is it bad that this person lived? It is of course bad, even tragic, that his life is cut short. But this is not the claim made by (2). According to (2), though this person's life is well worth living, it would have been better if he had never lived. This is not plausible. Consider next parenthood, one of the "major respects that generally make human lives . . . worth living." Is (2) plausible when applied to those who cannot have children, but have

9. p. 106.

10. p. 105.

11. footnote 23.



lives that are well worth living? Is it bad that such people ever live? No. Consider next a severe and lifelong handicap. Think of someone born blind whose life is well worth living. Setting aside effects on other people, is it bad that such a person ever lives? Once again, this is not plausible.

Consider next some handicap whose effects are more severe. Suppose that, because someone has such a handicap, it is *not* true that his life is *well* worth living. Here (2) is more plausible. There may be people whose lives, though worth living, are so diseased and deprived that, even apart from effects on others, it seems bad that these people ever live.

Consider, finally, slaves. Here again (2) is more plausible. We object to slavery even when the slaves have lives that are well worth living. Consider Athens as it actually was, with its wonderful culture supported by slaves. Compare this with Athens as it might have been: with the same free citizens, and the same culture, but with the work of all the slaves done by machines. According to (2), this would have been intrinsically better. Other things being equal, it would have been better if none of the slaves had existed, and only machines had existed in their place. We might here agree.

I have suggested that we might accept (2) when it applies either to slaves, or to people whose lives are not well worth living. If this is all that (2) covers, it does not solve the Future Individuals Paradox. It explains the wrongness of having the Handicapped Child only if this handicap is very severe. It cannot explain why the 14-Year-Old Girl ought to postpone having her child. Nor can it criticize our choice of the Risky Policy. If someone's life is well worth living, though he dies young, it is not intrinsically bad that this person ever lives. So (2) cannot cover the people killed in the catastrophe. The deaths of these people will be bad for others, such as their surviving relatives. But (2) can criticize these effects only if it is intrinsically bad that these surviving relatives ever live. Since their lives are on the whole well worth living, this is not plausible. Similar remarks apply to those whom the catastrophe disables.

These remarks point to a wider conclusion. (2) compares outcomes in which different numbers of people would be born. It therefore has the wrong form for the Future Individuals Paradox. This arises in

Same Number Choices: cases where, though our choice affects the identities of future people, it does not affect their number. Whether we choose the Safe or Risky Policy, or whether we Deplete or Conserve, we can permissibly assume that the same number of people will later be born. There would of course be *some* difference in the numbers. But this we can morally ignore, since it would either not now be predictable, or would not be the feature which is morally important. Consider our choice of the Risky Policy, which may later cause a catastrophe. We can *perhaps* predict that, if the catastrophe occurs, somewhat fewer people will ever be born. But this is not the objection to our choice. The objection is that, if we had chosen the Safe Policy, there would have been no such catastrophe. The problem is to explain this objection, given that our choice will be worse for no one. The problem arises because our choice affects the identities of future people. If we had chosen the Safe Policy, *different* people would have later lived, and been spared the catastrophe. We cannot explain this objection by appealing to the intrinsic badness of one set of possible future lives. We must compare two such sets: those of the different people who, on the different policies, would later live.

This is clearest in the case of Depletion. If we choose Depletion, this may greatly lower the quality of life in the further future. But those who will later live will not have "restricted" lives. I supposed, in my diagram, that their lives would be no worse than ours. The objection to our choice cannot be that, because of the intrinsic quality of these future lives, it would be bad if such lives are lived. The objection is that the people who will later live, though they will have good lives, will be much worse off than the different people who would have later lived if we had Conserved. Since it does not compare different possible lives, (2) cannot explain this objection.

I suggested another reading of Kavka's principle. According to (1), it is intrinsically worse for someone to live a restricted life than for someone else, in his place, to live an unrestricted life. This has the right form for Same Number Choices. But if we give the principle this form, there seems no need to use the word "restricted." We can claim

(3) If the same number of lives would be lived either way, it would be intrinsically worse if those who live are worse off than those who would have lived.

This seems intuitively plausible. And it provides objections to the choices in my four examples.

(3) may seem less plausible when applied to some actual person. Consider the Case of the Handicapped Child. Suppose that this child's life is, as we predicted, less worth living than the life of a normal child. According to (3), it would have been better if, instead of this child, his mother had later conceived a normal child. We may shrink from claiming, of some actual person, that it would have been better if he had not been conceived. If we accept (3), this is what we must claim. We cannot consistently make a claim but deny that same claim later. If in 1990 it *would be* better if this woman waits and has a normal child, then in 2020 it *would have been* better if she had waited and had a normal child. If we cannot accept the latter, we must retract the former. But I suggest that, on reflection, the latter claim is acceptable. I believe that, if *I* was this handicapped child, I could agree that it would have been better if my mother had conceived a normal child instead of me. This claim is not, as it was with (2), that my existence is intrinsically morally undesirable. The claim is merely that, if my mother had conceived a different child, that would have been better. And this need not imply that I ought rationally to regret that my mother had *me*, or that *she* ought rationally to regret this. If she loves me, her actual handicapped child, this is enough to block the claim that she is irrational if she does not have such regret.<sup>12</sup> Even when applied to some actual person, (3) seems to me acceptable.

If (3) is intuitively plausible, and provides objections in our four examples, does it solve the Future Individuals Paradox? Only superficially. (3) merely restates our intuitions. And it is restricted to Same Number Choices. We shall need some wider principle to cover Different Number Choices. We can hope that this wider principle will both imply and explain (3).

One such wider principle is Kavka's first proposal. Kavka believes that causing someone to exist can benefit this person. To state the full belief: If some choice is a necessary part of the cause of the existence of a person with a life worth living, this choice thereby

12. I follow Robert Adams, "Existence, Self-Interest, and the Problem of Evil." *Nous* 13 (1979).

benefits this person. Kavka then appeals to a version of the principle that we ought to do what would benefit people most. On this version, we must compare the benefits to the different people who, if we made different choices, would exist. This is Kavka's *maximizing principle*.<sup>13</sup>

This principle extends the use of the word *benefit*. When we claim to have benefited someone, we are usually taken to mean that some act of ours was the *chief* or *immediate* cause of some benefit received by this person. On the maximizing principle, we benefit someone if we make some choice which is either good for this person, or better for this person than the alternative, even when such a choice is a remote causal antecedent of the benefit received by this person. All that needs to be true is that, if we had chosen otherwise, this person would not have received this benefit.

We should next note that, in deciding what would benefit people most, we should count, for each possible act of ours, *all* of the benefits that would later be received if and only if we do this act. This may not be the natural view. Consider:

*The Joint Rescue Mission.* The lives of a hundred people are in danger. These people can be saved if five other people join in some rescue mission. If any of the five fails to join, no one will be saved. All do join.

How much good does each rescuer do? On one view, each produces his share of the total benefit. Since a hundred lives are saved, by five rescuers, each saves twenty lives. Call this *the Share of the Total View*. Opposed to this is *the Full Comparative View*. On this view, the good done by each person is the full difference that he makes, given what the others do. In our example, each of the five plays a necessary part in the saving of a hundred lives. It is true of each that, if he had not played his part, all of the hundred would have died. On the Full Comparative View, the benefit produced by each is the saving of a hundred lives.

Share of the Total Theorists think this absurd. If this is the benefit produced by each, must we not conclude that the five together save five hundred lives? But this is not true. In general, on the Full Com-

parative View, the sum of the benefits produced will not equal the sum of the benefits received. Only the Share of the Total View ensures that these two sums are equal.

Though natural, this objection has no force. These two sums need not be equal. The Full Comparative View does not answer the question of how great are the benefits which people receive. It answers the different question of which act, within some range of acts, would benefit people most. As an answer to *this* question, it is clearly the better view. Let us add one feature to our case. Suppose that I am one of the five who could jointly save a hundred lives. I also have some special skill by which, singlehandedly, I could save ninety of these lives. In order to do this, I must withdraw from the joint rescue mission, and the remaining ten people will die. On the Share of the Total View, this is what I ought to do. I would thereby benefit people more. I would save ninety lives. If instead I join the rescue mission, my share of the benefit produced is only the saving of twenty lives. I can therefore do more good if I withdraw from the mission. But this is clearly wrong. On this alternative, ten lives are needlessly lost. I ought to join the mission, so that all the hundred lives are saved. Only the Full Comparative View gives the right answer here. Only on this view can I claim that, if I join the rescue mission, the benefit that I produce is the saving not of ninety but of a hundred lives. We merely need to add that, in making such a claim, I do not imply that I *alone* produce this benefit. This answers the objection raised above.

In deciding which act would benefit people most, we should use the Full Comparative View. Suppose that I can do either P or Q. In deciding which would benefit people more, we should compare *all* of the benefits and losses that people would later receive if and only if I do P rather than Q, and *all* of the benefits and losses that people would later receive if and only if I do Q rather than P. The act which benefits people more is the one which, in this comparison, would be followed by the greater *net* sum of benefits—that is, the greater sum of benefits minus losses. It is irrelevant that many other acts will also be necessary causal antecedents for the receiving of these benefits. On Kavka's maximizing principle, we ought to do what, in the above sense, would benefit people most. And we include, among the benefits, that of receiving a life worth living. This benefit is greater if this life is more worth living.

This maximizing principle both implies and explains claim (3). Thus it explains why the mother of the Handicapped Child ought to have waited. Having a handicapped child benefits this child less than having a normal child would have benefited that child. There is a similar objection to our choice of Depletion. This choice will benefit those who later live, since their lives will be worth living, and they would not have existed if we had chosen Conservation. But our choice of Depletion benefits these people less than the choice of Conservation would have benefited those who would have later lived. Similar remarks apply to the other cases.

Should we accept this maximizing principle? Kavka claims that it needs to be revised, so that it does not “normally entail . . . a duty to reproduce.”<sup>14</sup> Since he does not say how his principle should be revised, I shall consider two suggestions.

One is that act utilitarianism is too demanding. On this view, we have no duty to maximize utility when this would require from us too great a sacrifice, or too great an interference in our lives. If some couple do not want to have a child, doing so might be such an interference. Such a couple would then have no duty to have children. If we revise Kavka’s principle by adding these claims, the principle may be part of a pluralist morality. We could add other principles, such as a principle of justice, or Kavka’s modified Kantian imperative.

Another common view is that, even if some child would have a life that is well worth living, this provides no moral reason to have this child. It provides no reason even when other things are equal. Kavka would reject this view. His maximizing principle essentially appeals to the benefits that people would receive if, because of our choice, they come into existence and have lives worth living. The principle’s distinctive claim is that such benefits do provide moral reasons. Kavka would not want his revised principle to deny this claim. He would prefer the revision that I proposed above.

Kavka claims that his principle needs a second revision. It must not imply that, if one decides to have a child, one ought “to produce the happiest child one could.”<sup>15</sup> As he writes, “we do not believe that

14. p. 99, where he writes, “We do not . . . feel that normal prospective parents would be under an obligation to procreate even if this would maximize social utility.”

15. p. 100.

a couple would be obligated to take genetic-enhancement pills if this would insure the production of a (different but) 'better' child—one that would be happier or contribute more to others' happiness. Such people might justifiably prefer to reproduce naturally, without such interference. Yet the maximizing principle would seem to imply the opposite."

Once again, Kavka does not state the needed revision. His words may suggest

(4) We would have no duty to have one child rather than another simply on the ground that the one child would be happier, or would have a life that would be more worth living.

But Kavka would reject (4). It denies his principle's distinctive claim. And he thinks it wrong to have a handicapped child when by merely waiting one could have a different child who would be normal. I expect he would also believe that the 14-Year-Old Girl ought to have her child later, so that she can give her child a better start in life. These beliefs are undermined by (4).

I suggest that Kavka's revised principle should claim

(5) We would have no duty to use artificial gene-enhancement.

This may seem too narrow, or suspiciously *ad hoc*. But artificial gene-enhancement threatens deep beliefs about equality, freedom, variety, and human dignity. We might be able to explain (5) by appealing to these beliefs. Nor is (5) too narrow. We shrink from gene-enhancement. But if we are going to have a child, and we know of other ways in which we could have a child who would be happier, or have a life that is more worth living, it seems plausible to claim that, if other things are equal, this is what we ought to try to achieve.

It may help to restate my proposed version of Kavka's principle. This has become

*The Revised Maximizing Principle:* If our choice is a necessary part of the cause of the existence of a person with a life worth living, our choice thereby benefits this person. Other things being equal, we ought to do what would benefit people most. In deciding how to do this, we must compare the benefits to the different people who, if we make different choices, would exist. We have no

duty to benefit others when this would require from us too great a sacrifice. Most of us therefore have no duty to have unwanted children. But if we do decide to have a child, and other things are equal, we ought to do so in the way that would benefit people most, unless this would involve artificial gene-enhancement, or—as before—require too great a sacrifice.<sup>16</sup>

This statement raises further questions. How should we decide, for instance, what in different cases would be too great a sacrifice? And how much weight should we give to this principle when other things are *not* equal—when this principle conflicts with one of our other moral principles? But we can ignore these questions here.

In this revised form, the maximizing principle still provides objections to the choices in my four examples. It therefore offers a solution to the Future Individuals Paradox.

### III. KAVKA'S REJECTION OF HIS FIRST PROPOSAL

Kavka rejects his maximizing principle, claiming that, when it is revised, it becomes inadequate. This part of Kavka's argument I find puzzling.

He makes his claim when discussing his *Case of the Slave Child*. Some couple have three alternatives: they could either have a child whom, in advance, they have sold into slavery, or have a different child who would be free, or have no child. Kavka claims that, according to his revised principle, having a slave child would be permissible. Since he believes that this act would be seriously wrong, he rejects his principle.

What does his principle here imply? One possibility, though unlikely, cannot be excluded. Having a slave child might produce such great benefits both to his parents and to the slaveholder that, on balance, this would be the act that would benefit people most. If this were so, the revised principle would imply that, if other things are equal, this is what the couple ought to do.

<sup>16</sup> In his footnote 13 Kavka leaves open the question of whether his principle should be concerned with total net benefits, or with average net benefits per person. I have assumed the former. He also suggests that he wants his principle to be part of a pluralist morality, to be combined with other principles such as those about just distribution. My proposed revision allows this.



Does this give us grounds to reject the principle? Only if it claims to cover the whole of morality. In the form that I proposed, it does not claim this. It covers the part of morality that is concerned with beneficence, or the promoting of people's interests. Since it includes the phrase "if other things are equal," the principle does not imply that having a slave child *must* be what this couple ought to do. We can both accept the principle and believe that having a slave child would be wrong. We could explain this by appealing to a principle of justice, or to Kavka's modified Kantian imperative, which forbids having a child as a mere means. The fact that we need such other principles, in some cases, does not show that we should reject the Revised Maximizing Principle. This principle may still give the right account of the part of morality concerned with beneficence. And it may, in particular, solve the Future Individuals Paradox.

We can now turn to the more likely version of Kavka's Case. This is the version where, compared with having a slave child, having a free child would benefit people more. This is the more likely version because a slave child's life would be likely to be much less worth living than a free child's life. Having a slave child would then benefit this child much less. On my proposed revision, the maximizing principle here implies that having a slave child would be wrong. The couple have no duty to have any child, but, if they do, they ought to have a free child.<sup>17</sup>

Kavka disagrees. And this disagreement has, we shall see, wider implications. Kavka writes: "(6) . . . the couple would not have an obligation to remain childless rather than to produce a slave child. (7) Nor would a duly restricted version of our maximizing principle imply that they are obligated to produce a nonslave child in preference to both alternatives. Hence (8), according to the principles we have so far considered, the deal with the slaveholder would be permissible."<sup>18</sup>

How does (8) follow from (6) and (7)? (7) denies that, on the

17. It may be objected that, if they cancel the deal with the slaveholder, they will lose his offered \$50,000. Can our revised principle require them to bear so great a sacrifice? We can reply that turning down a large possible gain does not count as such a sacrifice.

18. p. 100–101. (I have added the numerals.)

revised principle, the couple have a duty to have a free child in preference to both of the other alternatives. This does not imply that both of these alternatives would be permissible. (7) is true because one of these alternatives, remaining childless, is on the principle permissible. The other alternative, having a slave child, might be wrong. And this might be wrong even if (6) was also true—even if this couple have no duty to remain childless rather than have a slave child. They have no duty to remain childless because it would not be wrong to have a free child. And we can claim that, if they have a child, this is what they ought to do. These coherent claims are what, in the more likely version of this case, Kavka's revised principle implies.

As we have seen, Kavka disagrees. I have questioned the argument that he gives. Has he other arguments in mind? He does not appeal to the assumption that having a slave child would benefit people most. It might be thought that, for his second revision, he intends something like (4). But we have seen that he would reject (4). It is implausible, and unnecessary. The revision Kavka wants is sufficiently stated by (5) and the rest of my proposal.

Kavka may be appealing to the *transitivity principle* given in his footnote 16. This claims that, "if it would be permissible to do A if A and B were the alternatives, and would be permissible to do B if B and C were the alternatives, then it is permissible to do A if A, B, and C are the alternatives." By "alternatives" Kavka must here mean "only available alternatives." So interpreted, his transitivity principle cannot be generally assumed. It holds only within certain moral theories, such as act utilitarianism. What is permissible depends on what alternatives are available. Comparative permissibility therefore may not be a transitive relation when applied to cases where there are different alternatives available. Kavka's own views provide one example. He believes that, if one's only alternatives are to have a slave child or remain childless, it would be permissible to have a slave child.<sup>19</sup> And he believes that it would be permissible to remain childless rather than have a free child. On his transitivity principle, he would be forced to conclude that, even when one could have a free child, having a slave child would be permissible. This is

19. His footnote 24.

the very conclusion that he is most anxious to avoid. He can do so, since he can reject his transitivity principle.

It may help to give a new example. Suppose that I have three alternatives:

- A: at some great cost to myself, saving a stranger's right arm;
- B: doing nothing;
- C: at the same cost to myself, saving both the arms of this stranger.

If the cost were great enough, most of us would believe both (9) that, if only A and B were possible, neither would be wrong, and (10) that, if only B and C were possible, neither would be wrong. By Kavka's transitivity principle, if all three are possible, A must be permissible. It must be permissible to save only the right arm of this stranger, even though I know that at no extra cost to myself (or others) I could save both his arms. But would this be permissible? Saving the stranger's right arm at a great cost to myself may be admirably altruistic, but it would also be grossly perverse. If I am prepared to bear this cost, why do I not save *both* the stranger's arms? We might claim (11) that, if I bear the cost, and other things are equal, this is what I ought to do. If claims (9) to (11) are at least tenable, we can reject Kavka's transitivity principle.<sup>20</sup>

In the Case of the Slave Child, Kavka may instead be appealing to

20. Kavka might support his principle with this new claim: "If it would be permissible to do A if A and B were the only alternatives, A must be at least as good as B." Even when applied to cases where there are different alternatives available, *at least as good as* may be a transitive relation. On this new claim, if having a slave child is permissible when one's only alternative is remaining childless, the former must be at least as good as the latter. And if remaining childless is permissible when one's alternative is having a free child, the former must be at least as good as the latter. By transitivity, having a slave child must be at least as good as having a free child. We would again be forced to conclude that the deal with the slaveholder would be permissible.

Reconsider the case where, at great cost to myself, I could save a stranger's arms. We may here believe that doing nothing would be morally permissible. Does this imply that doing nothing would here be morally at least as good as saving the stranger's arms? This seems doubtful. We might believe that, while doing nothing would be morally permissible, it would be morally better to save the stranger's arms. If this belief is at least tenable, we can reject the new claim suggested above.

a more particular assumption. He may assume that, since having no child would benefit people less than having a slave child, the latter must be permissible if the former is. It may seem that an act must be permissible if it would benefit people more than some alternative that would be permissible. This is implied by certain moral theories, such as act utilitarianism. But it cannot be generally assumed. Reconsider my new example. We may claim (12) that it would be permissible here for me to do nothing. Saving one of the stranger's arms would benefit people more. Does it follow that it would be permissible to save only one of his arms, when I could just as easily save both? As I have suggested, we may claim (11) that, if I decide to bear the cost, and everything else is equal, I ought to save both arms rather than saving only one. If claims (12) and (11) are at least tenable, an act can be wrong even if it benefits people more than some permissible alternative. We can apply these remarks to having children. We may claim both that it is permissible to remain childless and that, if one does have children, one ought, other things being equal, to do so in the way that would benefit people most. This claim is implied by my revised form of Kavka's principle, and, in the more likely version of his case, this condemns having a slave child.

I conclude that, if it takes this form, Kavka's revised principle does not deserve his criticism. There may be cases where it does not explain our moral views. In such cases we must appeal to some other principle. But this is no objection to the revised principle if it is put forward, not as explaining the whole of morality, but as explaining beneficence, or the part of morality concerned with promoting people's interests. I have also claimed that, in one version of his case, the revised principle condemns the couple's choice.

This last claim has wider implications. Return to the Case of the Handicapped Child. This child's mother had three alternatives: to have a handicapped child, to have a different child who would be normal, and to have no child. This is like Kavka's Case with "handicapped" replacing "slave." If Kavka's revised principle could not criticize having a slave child, it could not criticize having the Handicapped Child. Nor could it tell the 14-Year-Old Girl to wait and have her child later. If the principle could not support these claims it could not solve the Future Individuals Paradox.

I have argued that it does support these claims. And it can criticize our choice both of the Risky Policy and of Depletion. It therefore offers a solution to the Paradox. Kavka's Revised Maximizing Principle, his first proposal, therefore seems to achieve more than he believes. In contrast, his other two proposals achieve less. In the less likely version of the Case of the Slave Child, we may need to appeal either to Kavka's principle about restricted lives, or to his modified Kantian imperative. And this last principle may state our main objection both to having the slave child, and to having a child merely so that he can be a kidney donor. But, as I have argued, these two principles do not solve the Future Individuals Paradox.

#### IV. WHETHER CAUSING SOMEONE TO EXIST CAN BENEFIT THIS PERSON

Does Kavka's maximizing principle really solve the Paradox? We can first reconsider one of its premises. This is the belief that, if some choice is a necessary part of the cause of the existence of a person with a life worth living, this choice thereby benefits this person. I shall try to show that this belief is not, as many claim, obviously mistaken.

Some objectors claim that life cannot be judged to be either better or worse than nonexistence. But life of a certain kind may be judged to be either good or bad, either worth living or worth not living. If a certain kind of life is good, it is better than nothing. If it is bad, it is worse than nothing. We should emphasize that, in judging that some person's life is worth living, or better than nothing, we are not claiming that it would have been worse for this person if he had never existed. Such judgments are most easily made about the last part of some life. Consider someone dying painfully, who has already made his farewells. This person may decide that what he has before him, if he lingers on, would be worse than nothing. If such claims can apply to parts of a life, they can perhaps apply to whole lives.

The objectors might now appeal to

*The Two-State Requirement:* We benefit someone only if we cause him to be better off than he would otherwise at that time have been.

This requirement seems too strong. It implies that saving someone's life cannot benefit this person, since the person saved is not better off than he would have been if he had ceased to exist. In the case of saving life, it seems defensible to relax the Two-State Requirement. We understand the special reason why, in this case, the requirement is not met. We might claim that, because of the special feature of the case, the requirement need not here be met. If the rest of someone's life would be worth living, we might count saving his life as a special case of benefiting him.

The objectors might now turn to

*The Full Comparative Requirement:* We benefit someone only if we do what will be better for him.

They could then say: "In causing someone to exist we cannot be doing what will be better for him. If we had not caused him to exist, this would not have been worse for him." Unlike the Two-State Requirement, this new requirement allows that saving someone's life can benefit this person. We can claim that it can be worse for someone if he dies, even though this does not make him worse off. (We would here be rejecting *the Lucretian Premise*: that some event can be bad for someone only if it makes him later suffer, or at least have regrets. Though not absurd, this can be rejected.)

Because it can cover saving life, the Full Comparative Requirement is more plausible than the Two-State Requirement. But if we can relax the latter, in the case of saving life, it may be defensible to relax the former, in the case of giving life. We can admit that, in every other kind of case, we benefit someone only if we do what will be better for him.<sup>21</sup> In the case of giving someone life, we understand the spe-

21. It may be objected that we can sometimes benefit people, even though what we are doing is not better for them. This can be true when our act, though sufficient to produce the benefit, is not necessary. Suppose that I could save either Ann's life or John's arm. I know that, if I do not save Ann's life, someone else certainly will; but no one else can save John's arm. If I save Ann's life, I will thereby benefit her. And I give her a greater benefit than the benefit I would give to John if I saved his arm. But, for moral purposes, this is not the way to judge benefits. In the case described, it would be clearly wrong for me to save Ann's life rather than John's arm. The benefit I give to Ann is not morally significant because, in giving her this benefit, I am not doing what will be better for her. This kind of case therefore poses no objection to the Full Com-

cial reason why the alternative would not be worse for him. We might claim that, in this special case, the requirement need not be met. Suppose we have allowed that saving someone's life can benefit this person. If my own life is worth living, it may then have benefited me to have had my life saved at any time after it started. Would it be plausible to claim that, while it benefited me to have had my life saved *just after* it started, it did not benefit me to have had it started?

Causing someone to exist is a special case because the alternative would not have been worse for this person. We may admit that, for this reason, causing someone to exist cannot be *better* for this person. But it may be *good* for this person.<sup>22</sup> In this move from "better" to "good," we admit that the Full Comparative Requirement is not met. But we would still make two kinds of comparison. If it can be good for some if he is caused to live, *how* good this is for this person will depend on how good his life is—how much his life is worth living. And we can make interpersonal comparisons. Suppose that Jack's life would be barely worth living, while Jill's life would be well worth living. We can then claim that, if we cause Jack to exist, this would be good for Jack, but it would be *less* good for Jack than causing Jill to exist would be for Jill.

We can next point out that these claims avoid a common objection. When we claim that it was good for someone that he was caused to exist, we do not imply that, if he had not been caused to exist, this would have been bad for him. And our claims apply only to people

---

parative Requirement. It is enough to make that Requirement read: We give someone a *morally significant* benefit only if we do what will be better for him. (Compare my earlier claims about the Full Comparative View. The amount of good that I do, if I choose P rather than Q, depends on the full difference that my choice will make. I should compare all and only the later benefits for which my choices would be *necessary* causal antecedents. I should ignore those benefits for which my choices are *sufficient* but *not* necessary antecedents.) These remarks about benefits can be applied to harms. On the Full Comparative Requirement, we impose on someone a morally significant harm only if we do what will be worse for him. This might be denied, and the denial held to explain the objection to our choice of the Risky Policy. Because it is often claimed that harms are not analogous to benefits, this denial cannot simply be dismissed. But I have no space to pursue the issue here. I shall merely remark that such a view about harms cannot provide a full solution to the Future Individuals Paradox. It cannot, for instance, criticize our choice of Depletion.

22. I owe this suggestion to Jefferson McMahan.

who are or would be actual. We make no claims about people who are or would remain merely possible. We are not claiming that it is bad for possible people if they do not become actual.

We might end with these remarks. We have considered three things: never existing, starting to exist, and ceasing to exist. We have suggested that, of these, starting to exist should be classed with ceasing to exist. Unlike never existing, starting to exist and ceasing to exist both happen to actual people. That is why, we might claim, they can be either good or bad for these people. The contrary claim is that starting to exist should be classed with never existing, and that neither can be either good or bad for people. The reason sometimes given is that, if we had not started to exist, we would never have existed. But we are not claiming that starting to exist can be either good or bad for people when it does *not* happen. Our claim is about starting to exist *when it happens*. We admit one difference between starting to exist and ceasing to exist. If it is good for someone if his life is saved, it would have been worse for him if he had died. Such entailments generally hold. For almost all events, if their occurrence would be good for people, their non-occurrence would have been worse for these people. But, we may suggest, there is one special event whose occurrence can be good for an actual person, even though its non-occurrence would not have been worse for this actual person. This event, unsurprisingly, is the coming-to-be-actual of this person.

These remarks are not conclusive. Further objections could be raised. My claim is only that, if we believe that causing to exist can benefit, we are not obviously mistaken. I shall consider later how, if we reject this belief, we might hope to solve the Future Individuals Paradox.

#### V. VARIETIES OF BENEFICENCE

If we believe that causing to exist can benefit, as Kavka does, we must decide between different versions of the principle of beneficence. It will help to define some more phrases. Suppose that we can do either P or Q. Call the people who will ever exist if we do P *the P-people*. Suppose that we choose P. Call our choice



“worse for people” in the *narrow sense* if the choice of P rather than Q would be either bad for, or worse for, the P-people,<sup>23</sup>

and

“worse for people” in the *wide sense* if the choice of P would be less good for the P-people than the choice of Q would be for the Q-people.

In Different Number Choices, “less good for” is ambiguous. Call our choice

“worse for people” in the *wide total sense* if the choice of P rather than Q would give to the P-people a smaller total net benefit than the benefit that the choice of Q rather than P would give to the Q-people,

and

“worse for people” in the *wide average sense* if the choice of P rather than Q would give to the P-people a smaller average net benefit per person than the choice of Q rather than P would give to the Q-people.

We can now distinguish three principles of beneficence. All claim that, if other things are equal, it is wrong knowingly to make some choice that would be worse for people than some other choice that we could have made. On the *Narrow Principle*, “worse for people” has its narrow sense. This principle condemns some choice only if this choice would be bad for, or be worse for, some of the people who would ever live. On the *Wide Total Principle*, “worse for people” has its wide total sense. This principle is very similar to Kavka’s Revised Maximizing Principle. To adapt Kavka’s words, it “requires weighing the potential benefits to different actual and possibly actual people, and choosing the act that produces the greatest total *net* benefit—that is, the greatest sum of benefits minus burdens.”<sup>24</sup> In the *Wide Average*

23. To avoid possible contradictions, we must add, “and by more than the amount, if any, by which the choice of Q rather than P would be either bad for, or worse for, the Q-people.”

24. p. 99.

*Principle*, we substitute the words “the greatest average net benefit per person.”

I have defended Kavka’s belief that causing to exist can benefit. But it is hard to decide whether we should accept this belief. If we do, we must decide between the principles we have just distinguished. This is another difficult decision.

Accepting either of the Wide Principles would have one advantage. Like Kavka’s similar principle, the Wide Total Principle offers a solution to the Future Individuals Paradox. So does the Wide Average Principle. The objectionable choices in our four examples would be “worse for people” in the two wide senses. In Same Number Choices, where the Paradox arises, the two Wide Principles coincide. In these cases these principles yield plausible conclusions. But this does not show that either principle is acceptable. We must ask what they imply about Different Number Choices. They may here yield conclusions that are not plausible.

Let us start with the simplest case:

*The Happy Child.* Some couple cannot decide whether to have another child. They can assume that, if they did, they would love this child, and his life would be well worth living. They have several reasons for wanting another child. But they also have several reasons for not wanting this, such as the interference it would bring to their careers. Like many others, this couple cannot decide between these two sets of conflicting reasons. They can also reasonably assume that, if they have this extra child, this would not predictably be on balance either better or worse either for them or for other people.

If causing to exist can benefit, having this child would benefit him. And it would not be predictably worse for anyone else. Choosing not to have this child would thus be “worse for people” in both of the wide senses. It would give to people a smaller total benefit, and a smaller average benefit. (The latter is true because failure to have the child gives to people *no* net benefit. Since having this child would give him a benefit, it would give to people a greater average benefit per person.) Both Wide Principles imply that, if other things are equal, this couple ought to have this child.

Can we assume that other things are equal? For this to be so, it must first be true that none of our other moral principles applies to this case. It seems that we can assume this. It must then be true that other things are equal according to our principle of beneficence. We can here distinguish four possibilities. Other things would be equal in the strongest sense if it was true that, whether or not this couple have the extra child, these alternatives would be *equally good* both for them and for everyone else. While this might be true of many choices, it is most unlikely here. A more likely possibility is that, of the couple's two alternatives, *neither would be worse than the other* for the couple and everyone else. This is a different possibility if there is only partial comparability, since *not worse than* does not then imply *at least as good as*. Two further possibilities are that neither alternative would be *predictably* worse for the couple and for other people, and that the couple *cannot decide* between their conflicting reasons.

What we would be justified in claiming depends on which of the above is true. Consider:

*The Doctor.* Some doctor can work in either England or India. When she thinks only of herself, she has several reasons for choosing England, and several reasons for choosing India. She cannot decide between these two sets of conflicting reasons. And neither choice would be predictably worse for herself. She also knows that, if she works in India, she will be able to save more lives.

That she could save more lives gives this doctor a moral reason to choose India. Other things are, in a weak sense, equal. We could here claim that, if she cannot decide between her other conflicting reasons, she ought to be swayed by this moral reason. She ought to choose India, where she could save more lives. But suppose that she now chooses England. Is she open to moral criticism? It would not be clear that she is. She may have now decided, for example, that, if she chooses India, this *would* be worse for her. If that is so, her choice of England might be morally permissible. She has no duty to save extra lives if this would require from her too great a sacrifice. It would be quite different if her two alternatives would be *equally good* for her. Suppose that in her hospital she could use either of two methods of treatment, one of which would save more lives. We could here be

confident that, since other things are equal in the strongest sense, she ought to use the method that would save more lives.

Return now to the Case of the Happy Child. Our couple's choice is like the Doctor's choice of where to work. Other things are equal only in the weaker senses that the couple cannot decide, and that neither choice would be predictably worse either for them or for other people. If the couple do decide not to have the extra child, it would not be clear that they are open to moral criticism. They may have now decided that they do not want to have this child, or that doing so would be worse for them. But there is a different question to which our answer can be clear. The fact that our Doctor could save more lives gives her a moral reason to choose India. Do our couple have a moral reason to have the extra child? On the Wide Principles, they do. Since this child's life would be well worth living, having this child would give him a great benefit. If the couple cannot decide between their other conflicting reasons, they ought to be swayed by this moral reason. Since other things are equal, they ought to have this child.

Can we accept these claims? If we can, we do not yet have grounds for resisting the Wide Principles. But, as we have said, there is another common view. On this view, the case is not like that of the Doctor. The fact that she could save more lives is a moral reason. But the fact that a child would have a good life is no moral reason to have this child. It is no reason even if everything else is equal.

If we hold this common view, we must reject the Wide Principles. We will then lose their solution to the Future Individuals Paradox. But this may still be the better alternative. Whatever we think in the Case of the Happy Child, the Wide Principles may be, in other cases, clearly unacceptable. And the Paradox may have a different solution.

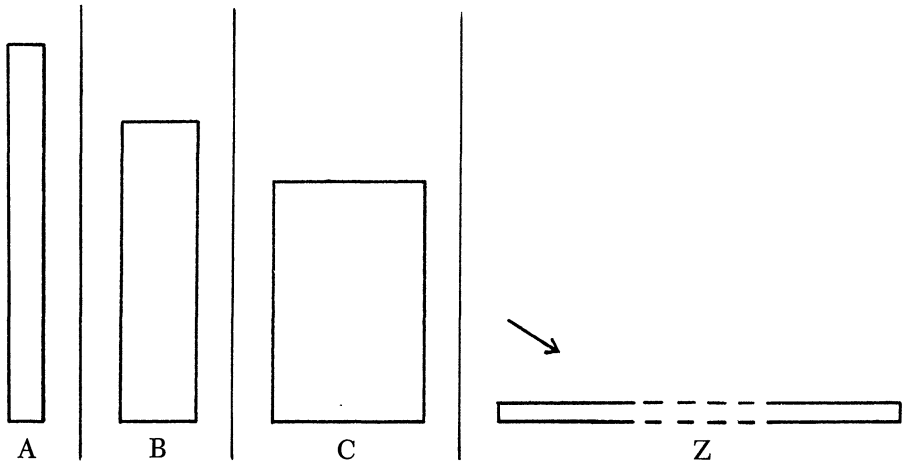
To help us choose between these alternatives, let us turn to cases on a larger scale.

## VI. THE REPUGNANT CONCLUSION

Let us consider various possible futures for one country, or mankind. I shall continue to talk of the quality of life. As before, we could think instead either of the level of happiness, or of the share per person of resources. We can assume that, in our examples, these three would

correlate. Our arguments would thus apply whichever we take to be most important.

Different outcomes may be represented thus :



The width of the blocks here shows the number of people living, the height shows their quality of life. I assume, for simplicity, that in these different outcomes there is neither social nor natural inequality. Thus in each outcome no one is worse off than anyone else. And no one would exist in more than one outcome.

In B there are twice as many people living as in A, and these people are all worse off than everyone in A. But the lives of those in B, compared with those in A, are *more than half as much* worth living. This claim need not assume precision. There may be only partial comparability. What the claim assumes is that a move from the level in A to that in B would be a decline in the quality of life, but that it would take much more than a similarly great decline before people's lives ceased to be worth living.

If B comes about rather than A, this would be "better for people" in the wide total sense. B is *in toto* better for the B-people than A is for the A-people. B is of course less good for *each* of the B-people than A is for each A-person. But since each B-person would benefit more than half as much as each A-person, and there are twice as many B-people, they *together* would benefit more. The Wide Total Principle

therefore implies that, if other things are equal, we ought to bring about B rather than A. Let us now extend this principle. It can claim that, if one of two outcomes would be “better for people” in the wide total sense, and other things are equal, this would be the better outcome. The principle now implies that, if other things are equal, B would be better than A.

By the same reasoning, Z could be best. Z is some enormous population whose members have lives that are not much above the level where life ceases to be worth living. A life could be like this either because it has enough ecstasies to make its agonies worth enduring, or because it is uniformly of poor quality. Let us imagine the lives in Z to be of this second drabber kind. If Z comes about, each of the Z-people would thereby benefit very little. But, if Z is large enough, they together would benefit most. The Wide Total Principle thus implies

*The Repugnant Conclusion:* For any possible and large population, say of eight billion,<sup>25</sup> all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, and be what we ought to bring about, even though its members have lives that are barely worth living.

As my choice of name implies, I find this conclusion hard to believe.

Some claim that we can ignore this conclusion, since other things would never be equal. But this seems doubtful. It is not clear that the coming-about of Z must infringe some other plausible moral principle. Thus it is not clear that it must involve either the violation of rights, or injustice. The only kind of injustice that could be involved is the breach of a principle about justice between generations. But this is irrelevant to our question in its purest form. We are asking whether, if Z comes about, this would be better than if A comes about. We could imagine a history in which only Z-like outcomes occur. The people in Z would then be no worse off than anyone who ever lives.

25. Many believe that, if the number of people who ever lived would be extremely small, there would be some value in the existence of extra people who had lives worth living. As the numbers increase, the value in the increase steadily declines, and reaches zero. This is why I write “eight billion.”

If we believe that Z would be worse than A, this could not here be because Z's occurrence would involve injustice.

If we wish to avoid the Repugnant Conclusion, we should not try to do so by appeal to principles covering some different part of morality. The conclusion seems *intrinsically* repugnant. It is an answer to the central question raised by Different Number Choices: whether, if the quality of life is lower, this can be made up for by its greater quantity. On the Wide Total Principle, the answer is yes. This principle implies that, provided that lives remain worth living, any loss in the quality of life could be made up for by a sufficient gain in its quantity. If we cannot accept the Repugnant Conclusion, this is what we must deny. We take a different view, in Different Number Choices, about the part of morality that is concerned with people's interests, the quality of life, and human well-being. If we think that Z is worse than A, our belief is this: the fact that people are so much worse off, even though their lives are worth living, cannot be made up for by the fact that there are so many more people living.

What of the comparison between A and B? Many people here believe that B would be worse than A. They believe that it would be worse for there to be more people if they would *all* be worse off. On the Wide Total Principle, as now extended, B would be better than A, C would be even better, and so on. If we cannot believe this, we have further grounds for resisting the principle.

Are these good grounds? Do these simplified examples provide a fair test for our principles? I believe that they do. A and B are acceptable simplifications of what would in practice be real alternatives. In any possible future there would in fact be some inequality. But it cannot distort our reasoning, on the central question I have raised, if we imagine the simpler case where this would not be so. A and B then represent two possible futures, for some country or mankind, given two slightly different rates of population growth over some period like a century. If we do not believe that B would be better than A, we have a practical test of the Wide Total Principle.

A and Z would not in practice be real alternatives. Some claim that, because of this, we need not try to avoid the Repugnant Conclusion. They might say: "Since this conclusion does not involve a

possible choice, it can be ignored. We need not test our principles in cases that could not occur.”

We can distinguish two kinds of impossibility. Call these *deep* and *technical*. An imagined case is deeply impossible if it requires a major change in the laws of nature, including the laws of human nature. There are two grounds for challenging cases that are deeply impossible. We may be unable to imagine what such cases would involve. And some would claim that our moral principles only need to be acceptable in the real world.<sup>26</sup>

It may help to remember here Nozick's imagined *Utility Monsters*. These are people who “get enormously greater gains in utility from any sacrifice of others than these others lose.”<sup>27</sup> Such an imagined person provides an objection to act utilitarianism, which “seems to require that we all be sacrificed in the monster's maw, in order to increase total utility.” As described by Nozick, such a person is a deep impossibility. Let us imagine the wretchedness of the world's population if all but this one person are denied anything above starvation rations, and all other resources go to this one person. We are then asked to suppose that this one person would be *so* happy, or have a life of *such* high quality, that this is the distribution which yields the greatest sum of happiness, or its equivalent in terms of worthwhile life lived. For this to be so, given the millions left in wretchedness, which could be so strikingly relieved by a small fraction of this one person's vast resources, this person's quality of life must, it seems, be millions of times as high as that of anyone we know. The qualitative gap between his life and ours must resemble the gap between ours, at its best, and the life of those creatures who are barely conscious—such as, if they *are* conscious, Plato's “contented oysters.”<sup>28</sup> It seems a fair reply that we cannot imagine, even in the dimmest way, the life of this Utility Monster. And this casts doubt on the force of the example. Act Utilitarians might say that, if we really could imagine

26. See, for example, R.M. Hare's discussion of the different levels of moral reasoning in his “Ethical Theory and Utilitarianism,” in *Contemporary British Philosophy*, H.D. Lewis, ed. (London, 1976), and his *Moral Thinking* (Oxford: Oxford University Press, 1982).

27. Robert Nozick, *Anarchy, State and Utopia* (New York: Basic Books, 1974), p. 41.

28. *The Philebus*, 21 c-d.



what such a life would be like, we might not find Nozick's objection convincing. His "Monster" seems to be a godlike being. In the imagined presence of such a being, our belief in our right to equality with him may begin to waver—just as we do not believe that the lower animals have rights to equality with us.

This reply has some force. But even a deep impossibility may provide a partial test for our moral principles. We cannot simply ignore imagined cases. Let us now return to my imagined Z. This imagined population is another Utility Monster. The difference is that the greater sum of happiness comes from a vast increase, not in the quality of one person's life, but in the number of lives lived. And *my* Utility Monster is neither deeply impossible, nor something we cannot imagine. We can imagine what it would be for someone's life to be barely worth living. And we can imagine what it would be for there to be many people with such lives. In order to imagine Z, we merely have to imagine that there would be *very* many. This we can do. So the example cannot be questioned as one that we can hardly understand. It may be true that we could not in practice face a choice between A and Z. Given some roughly finite stock of resources, we could not in fact produce the greatest sum of happiness, or its equivalent in terms of worthwhile life lived, by producing an enormous population whose lives are barely worth living.<sup>29</sup> But this would be merely technically impossible. In order to suppose it possible, we only need to add some assumptions about the nature and availability of resources. If it would be merely technically impossible to face a choice between A and Z, this does not weaken the comparison as a test for our principles. Different Number Choices raise the question whether loss in the quality of life could be made up for by a sufficient gain in its quantity. This is the question posed most clearly by comparing A

29. On some versions of *the Law of Diminishing Marginal Utility*, this is just what is implied. On these versions, each unit of resources produces more utility if it is given to people who are worse off, so that the most productive distribution will be one where everyone's life is barely worth living. There is here an obvious oversight. Large amounts of resources are needed to make each person's life even reach the level where life begins to be worth living. Such resources do not help to produce the greatest causally possible net sum of utility, when they are merely used to prevent *extra* people having lives that are worth *not* living (or have net disutility).

and Z. If we are convinced that Z is worse than A, we have strong grounds for resisting principles which imply that Z is better. So we have strong grounds for resisting the Wide Total Principle. We must try to show that, even when it is put forward merely as part of a pluralist morality, this principle gives the wrong account of beneficence.

An obvious move is to appeal instead to the Wide Average Principle. (As we shall later see, this is quite different from what some writers call the "Utilitarian Average Principle.") If what comes about is Z rather than A, this would be "worse for people" in the wide average sense. Causing Z rather than A would give to the Z-people a smaller average benefit per person than causing A rather than Z would give to the A-people. But this does not solve our problem. The Wide Average Principle does not *directly* imply the Repugnant Conclusion. But it can do so indirectly.

Suppose that we have brought about A. We might now face a new choice. Suppose that in a short time we could change A into B. It might be "better for people" in the wide average sense if we made this change. This would involve adding to the existing population as many new people. And the previously existing half would suffer a decline in their quality of life. So this change would be worse for them. But the change would bring a greater benefit to the newly existing half. Changing from A to B would therefore give to the B-people an average net benefit per person. And it might give them a greater average benefit per person than keeping A rather than changing to B would give to the A-people.<sup>30</sup> The Wide Average Principle would then imply that, if other things are equal, we ought to move from A to B. By similar reasoning, we ought then to move from B to C, and ought then to move from C to D. Z is the population which, in the end, we

30. To make the point more clearly, let us allow ourselves to assume precision. Let the level in A be 100, and the level in B 80. In the change from A to B the previously existing half each lose 20, and the newly existing half each gain 80. So the change from A to B gives to the B-people an average net benefit per person of  $\frac{80 - 20}{2}$ , or 30. Keeping A rather than changing to B would give to each of the A-people a benefit of 20. So changing from A to B would give to the B-people a greater average benefit per person than not changing from A to B would give to the A-people. (I discuss an objection to this reasoning at the end of footnote 36.)

ought to bring about. The Wide Average Principle could thus indirectly imply one part of the Repugnant Conclusion. We turned from the Total to the Average Principle hoping to avoid this conclusion. We have found that, if we are to do so, we must reject both the Wide Principles.

Let us now review the argument so far. In Part I we described what Kavka calls the Paradox of Future Individuals. There seemed to be moral objections to our choice of the Risky Policy, or of Depletion, or to the choice of having a handicapped or disadvantaged child, when by merely waiting one could have a child free from these burdens. The problem was to explain these objections, given our knowledge that these choices would be worse for no one.

In Part II we considered Kavka's three proposed solutions. It seemed that only his first proposal, the Revised Maximizing Principle, could provide a full solution. We then asked whether this principle provided an acceptable solution. In Part III it seemed to survive Kavka's own criticisms. In Part IV we discussed objections to the belief required by the principle: the belief that causing to exist can benefit. I claimed that this belief is not obviously mistaken. Though controversial, it may form part of an acceptable solution.

In Part V we saw that, if we believe that causing to exist can benefit, Kavka's Revised Maximizing Principle is not the only form of the principle of beneficence. Kavka's principle is very similar to my Wide Total Principle. But we could appeal instead either to the Wide Average Principle, or to the Narrow Principle. We had one ground for choosing one of the Wide Principles. Only these offered a solution to the Future Individuals Paradox. But this did not show that we should accept one of these principles. Both may be unacceptable when applied to Different Number Choices.

In asking whether this is so, we began with the simplest case, that of the Happy Child. If we believe that causing to exist benefit, and accept either Wide Principle, we must agree that our couple have a moral reason to have this child. If other things are equal, this is what they ought to do. If we could accept this conclusion, we did not yet have grounds to resist the Wide Principles. We have now turned to another range of cases, that of the possible states of the world from A to Z. Even if we agreed that our couple had a moral reason to have

the Happy Child, we may be unable to believe that B would be better than A, or that we ought to change A into B. But this is what the Wide Principles here imply. And we may find it even harder to accept their other implication, the Repugnant Conclusion. I shall assume that most of us wish to avoid this conclusion.

If we are to do so, we must reject both Wide Principles. This is done by those who, in the Case of the Happy Child, claim that our couple have no moral reason to have this child. We now have stronger motives to explore this common view. We can first ask whether it survives a well-known test.

## VII. THE ASYMMETRY

Consider:

*The Wretched Child.* Some woman knows that, if she has a child, he will be so multiply diseased that his life will be worse than nothing. He will never develop, will live for only a few years, and will suffer pain that cannot be relieved.

It seems clear that it would be wrong knowingly to conceive such a child. Nor would the wrongness primarily lie in the effects on others. The wrongness primarily lies in the predictably appalling quality of this child's life.

Suppose we took the common view that our couple have no moral reason to have the Happy Child. We now believe that it would be wrong knowingly to have the Wretched Child. On our view, it would be wrong to have a child whose life would be worth *not* living, but, even if other things are equal, it would be in no way wrong *not* to have a child whose life would be worth living. These two claims have been called *the Asymmetry*.<sup>31</sup>

How could we explain this Asymmetry? We might deny that causing someone to exist can be either good or bad for this person. But this could explain only half the Asymmetry: why our couple have

31. By Jefferson McMahan, in his "Problems of Population Theory," a Critical Notice of *Obligations to Future Generations*, ed. R.I. Sikora and Brian Barry, in *Ethics* (October 1981). This Critical Notice discusses many of the questions that I discuss here. In writing these two papers McMahan and I are greatly indebted to each other.

no moral reason to have the Happy Child. It would imply that, in having the Wretched Child, his parents cannot be doing something that is bad for him. So what should our objection here be? If having this child cannot be bad for him, our primary objection must it seems appeal directly to his unrelieved suffering. We must appeal to

*The Impersonal Misery Principle:* Other things being equal, we ought not to increase the sum of suffering.

It may be hard to accept this principle but reject

*The Impersonal Happiness Principle:* Other things being equal, we ought to increase the net sum of happiness.

But if we accept this last principle we still cannot explain the Asymmetry. This principle implies that our couple have a moral reason to have the Happy Child. Even if having this child cannot be good for him, it would increase the sum of happiness. It may be objected that suffering and happiness are morally dissimilar. Our moral reasons to prevent the former far outweigh our moral reasons to promote the latter. But this cannot explain the Asymmetry unless we have *no* moral reason to promote happiness. And, if we have accepted the Impersonal Misery Principle, it seems implausible to reject entirely its analogue for happiness.<sup>32</sup>

If we are to defend the Asymmetry, we must reject these two Impersonal Principles. We might claim that they take the wrong form, treating people as the mere *containers* of value. This *Milk Production Model* may be held to distort morality. We may thus appeal to

*The Person-affecting Restriction:* This part of morality, the part concerned with human well-being, should be explained entirely in terms of what is good or bad for those whom our acts affect.

This is the argument advanced by Narveson.<sup>33</sup> If we have denied that having the Happy Child can be good for him, we can now revive the claim that our couple have no moral reason to have this child. But,

32. See, for example, James Griffin's "Is Unhappiness Morally More Important Than Happiness?" *Philosophical Quarterly* 29 (January 1979).

33. In his pioneering "Utilitarianism and New Generations," *Mind* 76 (January 1967), to which much of the recent debate owes its existence. See also Jan Narveson's "Moral Problems of Population," *The Monist* 57, no. 1 (1973).

as we saw, this undermines the other half of the Asymmetry. We must admit that having the Wretched Child cannot be bad for him, and we have disallowed appeals to the Impersonal Misery Principle.

If we deny that causing someone to exist can be either good or bad for this person, the Asymmetry seems hard to explain. There is an alternative. We could retain the Person-affecting Restriction, claim that causing someone to exist can be either good or bad for him, and appeal to the Narrow Principle. The Asymmetry can now be fully explained. According to the Narrow Principle, it is wrong, if other things are equal, to do what would be either bad for, or worse for, the people who ever live. It is therefore wrong to have the Wretched Child. Since his life is worse than nothing, having this child is bad for him. But it is in no way wrong to fail to have the Happy Child, whose life would be well worth living. True if the couple had this child, this would be good for him. But if they do not have this child this would not be bad for him. And, in the case described, it would not be bad for anyone. This is why they have no moral reason to have this child.

This seems the better explanation of the Asymmetry. We should note that the Narrow Principle does not involve the familiar claim that our obligation not to harm is stronger than our obligation to benefit. We may wish to add this claim to the Narrow Principle of Beneficence—to add, as Ross did, some stronger principle about Non-Maleficence.<sup>34</sup> But the Narrow Principle makes no such distinction. If there will be someone to whom we have failed to give some possible benefit, our failure to do so will be worse for this person. If other things are equal, we have acted wrongly according to the Narrow Principle. We should also note that the distinction between the Wide and Narrow Principles is not the same as that between principles which do and do not appeal only to effects on those who ever live. Suppose that we refrain from conceiving the Wretched Child. We have acted rightly, according to the Narrow Principle. This is because, if we had conceived the Wretched Child, this would have been bad for him. In explaining why we acted rightly we appeal to a possible effect on someone who might have lived. (But, as we said, the effect would not have

34. Sir David Ross, *The Right and the Good* (Oxford: Oxford University Press, 1930), p. 21.

been on someone who remained merely possible. We never appeal to such effects. The effect would have been on an actual person.)

The distinction between the Wide and Narrow Principles is not, then, reducible to these other more familiar distinctions. It is a new distinction, opened up by the belief that, in causing someone to exist, we can thereby benefit this person. This belief breaks the ordinary entailment that, if some event would be good for people, this event's non-occurrence would be worse for people. With this entailment broken, the Wide and Narrow Principles diverge.

#### VIII. INTERIM CONCLUSIONS

Let us now take stock of our position. In the Case of the Happy Child, many people claim that our couple have no moral reason to have this child. When we considered population size, we were drawn to a similar view. We wish to avoid the Repugnant Conclusion, and, though perhaps less urgently, the claim that B is better than A. The simplest way to do this is to appeal, on a larger scale, to this common view about the Case of the Happy Child. We might claim that there is no moral reason, even if other things are equal, to cause extra people to exist who will have lives worth living.

There were two ways to defend this common view in the Case of the Happy Child. Both appealed to the Person-affecting Restriction. We could then either reject the belief that causing to exist can benefit, or instead accept this belief but appeal to the Narrow rather than either of the Wide Principles. Of these defenses of this view, the first cannot handle the Case of the Wretched Child. But the second defense covers this case. It therefore seemed the better explanation of the Asymmetry.

The problem with both defenses is that we lose our solution to the Future Individuals Paradox. If we believe that causing to exist can benefit, and appeal to either of the Wide Principles, we solve that Paradox. These principles both explain the objection to the choices we described in Part I. The objection is that these choices are in the two wide senses "worse for people." Though they benefit those who later live, they benefit these people less than the alternative would have benefited those who would have later lived. We cannot make

such claims if we either deny that causing to exist can benefit, or appeal to the Narrow Principle.

Can we solve the Paradox in a different way? Not if we continue to appeal to the Person-affecting Restriction. If we appeal to this Restriction, we believe that the part of morality concerned with well-being should be explained entirely in terms of what would be good or bad for those people whom our acts affect. We cannot then criticize the choices in Part I except with the claim that, in one of the wide senses, they are worse for people. So if we reject the Wide Principles we cannot solve the Paradox.

If we want another solution, we must abandon the Person-affecting Restriction. And we may have a second ground for doing so. We may not believe that, in causing someone to exist, we can thereby benefit this person. I argued that this belief was not obviously mistaken. But, after considering the arguments, we may reject this belief. If we do, it is again clear that, to solve the Future Individuals Paradox, we must abandon the Person-affecting Restriction. We must appeal to principles which are about well-being, or the quality of life, but are not just about what is good or bad for those whom our acts affect.

If we allow such principles, we must reconsider the two Impersonal Principles given above. On the Impersonal Misery Principle, it is wrong, if other things are equal, to increase the sum of suffering. On the Impersonal Happiness Principle, it is wrong, if other things are equal, not to increase the net sum of happiness. The second of these principles implies the first, and is the hedonistic form of the *Impersonal Total Principle of Beneficence*.

This principle offers a second solution to the Future Individuals Paradox, since it provides objections to the choices in Part I. But it clearly implies that B would be better than A, and that Z might be best. A greater total net sum of happiness might be found in a vast population, whose lives are barely worth living, just as a greater mass of milk might be found in a vast heap of bottles, each containing only a single drop.

Since it implies the Repugnant Conclusion, most of us would wish to reject the Impersonal Total Principle, whether it is phrased in terms of happiness or in terms of the quality of life. But, if we are



not appealing to the Person-affecting Restriction, it becomes unclear on what grounds we should criticize this principle, except that we reject what it implies.

We might simply claim that, while it is better if actual people become happier, or have a higher quality of life, it is not better if an extra life is lived which is happy, or worth living. Call this the claim that *extra lives cannot have intrinsic moral value*. Remember the central question raised by Different Number Choices: if those who live have a lower quality of life, can this be made up for by a sufficient increase in the number of lives lived? If there is *some* intrinsic value in each extra life lived that is worth living, this provides a natural ground for the answer yes. Even if the quality of life is lower, *enough* such extra lives would, it seems, have enough compensating value. We might reply that there are two kinds of value, which are not exchangeable. And this seems not implausible when applied to lives of very different quality. It may be our view when we compare A with Z. But we can reach Z indirectly, through B, C, and so on. To avoid the Repugnant Conclusion—the claim that Z is best—we must either deny that B is better than A, or make such a denial about two similar adjacent outcomes later in the series. The lives of those in B need not be of a quite different quality from the lives of those in A. It does not seem plausible to claim that, while there is indeed intrinsic value in each extra life lived at level B, *no* amount of *this* value could be equivalent to the value of the lives lived at level A. Perhaps we shall be forced to some variant of this view. But it is tempting to make a simpler reply: to claim that there is *no* intrinsic moral value in the existence of an extra person with a life worth living. Causing such a person to exist may, of course, on other grounds be what we ought to do. Perhaps, for example, the birth of this extra person would greatly benefit other people. (Thus medieval kings had duties to have children to prevent wars of succession.) But, if other things are equal, we have no moral reason to cause such an extra person to exist.

Another reply, less simple, might be built around the claim just made. We might claim that suffering is intrinsically bad, and that certain kinds of life are intrinsically good. Call such lives *wonderful*. What we deny to have intrinsic moral value are lives that are worth

living, but are below the Wonderful Level.<sup>35</sup> (When we claim that wonderful lives have intrinsic moral value, we are not claiming that these are morally good lives. Rather we are claiming that, whenever such a life is lived, this is an outcome which is good in the sense which has moral relevance.) If the lives in A are wonderful, this would explain why Z is worse than A. If we adopt this view, we could of course admit that, once someone exists, it is bad if things go worse for him. This can be so even if his life is in the intrinsically valueless band. We might thus appeal to the Narrow Person-affecting Principle of Beneficence, but add to this our two impersonal claims about the intrinsic badness of suffering, and of wretched lives, and the intrinsic goodness of wonderful lives.

By denying the Impersonal Total Principle, we have avoided the Repugnant Conclusion. But the problem, once again, is that we have no solution to the Future Individuals Paradox. Reconsider our choice of Depletion. Suppose that, even if we choose Conservation, the quality of life would not rise to the Wonderful Level. All of the different possible lives would be within the intrinsically valueless band. We may still believe that if, for some trivial benefit to ourselves, we greatly lowered the quality of life in the further future, there is at least some moral objection to our choice. But this cannot be explained either by our Narrow Principle of Beneficence, or by our new principles about impersonal value.

We therefore need at least one further principle. We can remember here claim (3): that, if the same number of lives would be lived either way, it would be worse if those who live are worse off than those who would have lived. This provides an objection to our choice of Depletion, and to the other choices described in Part I. I denied that (3) was a full solution to the Future Individuals Paradox, both because it seemed to need further explanation, and because it was restricted to Same Number Choices. We need some wider principle, covering Different Number Choices, which we find acceptable in all these choices, and which implies (3). Call this needed principle (X).

35. Though we deny that such a life has impersonal moral value, it will have *personal* value—value for the person whose life it is. Compare the distinctions drawn by Thomas Nagel in “The Limits of Objectivity,” *The Tanner Lectures of Human Values*, vol. 1 (Salt Lake City: University of Utah Press, 1980), pp. 121–26. My suggested view needs to be revised to include his distinctions.

One candidate for (X) is the *Impersonal Average Principle*. This is quite different from the Wide Average Principle discussed above. That was one version of a person-affecting principle, concerned with what is good or bad for those whom our acts affect. The Impersonal Average Principle does not appeal to such effects. According to this principle it is best, other things being equal, if there is the greatest average net sum of happiness per life lived, or if on average the lives that are lived go as well as possible, or have the highest average quality. This principle is accepted by many economists. Some make it true by definition.<sup>36</sup> I shall merely suggest below why I am sure that this prin-

36. See, for example, Paul Samuelson's *Economics*, 8th ed. (New York: McGraw-Hill, 1970), p. 551. I state the Impersonal Average Principle in a temporally neutral form. Some have stated the principle so that it is concerned only with the average of those living after we have acted. Such a principle tells us that it would be intrinsically better if we kill off all but the most ecstatic. But a temporally neutral Average Principle avoids this absurd conclusion. If we kill anyone whose life is worth living, we thereby lower the average per life lived.

The rest of this footnote will explain how the Impersonal Average and Total Principles differ from their Person-affecting counterparts, and defend the latter from the charge of double-counting. Those uninterested should not read on. To save words, I shall discuss only the hedonistic forms of these principles.

In the Case of the Happy Child, the Impersonal Average Principle tells our couple to have this child if, but only if, this will raise the average net sum of happiness per life lived. This will be so, in this case, if this child's net sum of happiness will be greater than this average net sum. The Wide Average Principle is not concerned with *this* average. As a Person-affecting Principle, it is concerned with the average benefit per person brought about by doing one of two things. If causing to exist cannot benefit, this principle does *not* tell our couple to have the Happy Child, even if doing so would raise the average happiness per life lived. If causing to exist can benefit, this principle *does* tell our couple to have this child, even if doing so would lower the average happiness per life lived. Having this child will benefit him. Failing to have this child will benefit no one. Of the two, having this child therefore gives to people a greater average benefit. This will be so even if this child would have less than the average sum of happiness. This is why the Wide Average Principle can, perhaps surprisingly, indirectly imply the Repugnant Conclusion.

We can next compare the Impersonal and Person-affecting Total Principles. If causing to exist cannot benefit, these will often differ. Thus the first but not the second would tell our couple to have the Happy Child. If causing to exist can benefit, the two principles would often coincide. They would both tell our couple to have the Happy Child. And they would coincide whenever we are comparing possible outcomes of which it is true that no one would exist in more than one of the outcomes. In such cases both principles directly imply the Repugnant Conclusion. But consider this example (where, for convenience,

ciple ought to be rejected. But there is another principle, of which this Impersonal Average Principle may be a misguided generalization. Reconsider A and B. We might claim

*The Two-Level Quality Principle:* Outcome B would be worse than outcome A if *all* of the B-people would be worse off than *all* of the A-people.

This seems plausible. And it would judge B to be worse than A even if the lives in A are not wonderful. (Some economists may start with such a claim and then move, rashly, from *all* to *on average*.)

---

we shall again assume precision). Eve already exists. We can either (1) leave her alone, with the result that the sum of happiness in her life will be 100, or (2) create Adam, with the result that both he and she will have sums of happiness of only 51. On the Impersonal Total Principle, (2) is better than (1). The total sum of happiness will be greater. The Wide Person-affecting Total Principle gives a different answer. On this principle, we ask which act would give a greater total net benefit to the people who will exist, if we do this act. If we do (1) rather than (2), Eve will gain 49, and no one will lose. The total benefit to those who will exist—that is, to Eve—will be 49. If instead we do (2) rather than (1), Adam will gain 51, but Eve will lose 49. The total net benefit to those who will exist will be only 2.

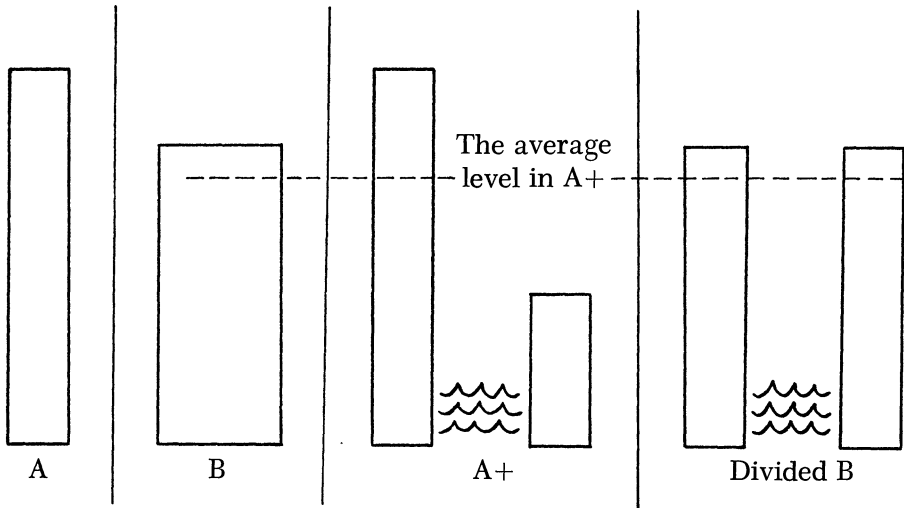
It may be objected that my Wide Principles involve double-counting. They do involve two calculations, where we usually need only one. In most of our moral thinking, we are concerned with cases where the same people will exist whatever we do. In such cases it is enough to ask what the net benefit would be if we do (a) rather than (b). We need not then ask what the net benefit would be if we do (b) rather than (a). Both calculations would yield the same moral conclusion. But in Different Number Choices we need both calculations. Take our case of Eve and Adam. In applying the Wide Total Principle to this case, the difference to Eve is counted twice. She counts as losing 49 if we create Adam, and as gaining 49 if we don't. But this is how we need to reason if we are comparing the total net benefits to those who will exist, given what we do. If we create Adam, his gain will slightly outweigh Eve's loss. But if we do not create Adam, Eve's gain will not be outweighed by anyone's loss. There will be no Adam to count as a loser. The total net benefit to those who will exist—that is, to Eve—will thus be much greater. This calculation favors Eve. But this is because, unlike her, Adam does not exist in both outcomes. Eve is favored in the way that we should expect from a Person-affecting Principle, even when it is concerned with the total sum of benefits. Similar remarks apply to the Wide Average Principle. (I have suppressed one problem. In their extended forms, which judge one of two outcomes to be better, both my Wide Principles can in certain cases lead to inconsistent judgments. We could avoid this problem by giving the principles a more complicated form. I believe that this would not affect my main conclusions.)

We have now reached these provisional conclusions. We wish to reject both of the Wide Person-affecting Principles, since they can imply the Repugnant Conclusion. We therefore admit that, to solve the Future Individuals Paradox, we must abandon the Person-affecting Restriction. We need some principle which is not about what will be good or bad for those whom our acts affect. We need such a principle whether or not we believe that causing to exist can benefit, for if we do believe this we will appeal to the Narrow rather than either of the Wide Principles, and this Narrow Principle cannot solve the Paradox. In Same Number Choices, where the Paradox arises, we accept claim (3). We need some broader principle, covering Different Number Choices, which will imply and explain (3). One broader principle is the Two-Level Quality Principle. This is compatible with (3). But it covers few cases. It is seldom true, of two possible populations, that *all* of one would be worse off than *all* of the other. We need some principle that covers every kind of case. I have suggested that one candidate, the Impersonal Average Principle, will prove unacceptable. And we wish to reject the Impersonal Total Principle, which in its hedonistic version judges an outcome to be better if it involves a greater net sum of happiness. We wish to reject this principle since it also implies the Repugnant Conclusion. But I have suggested that we might accept two parts of the Impersonal Total Principle. We might agree that it is intrinsically worse if there is more suffering, and that it is intrinsically better if more lives are lived above some Wonderful Level. We may add to this the Narrow Person-affecting Principle, and other principles, such as a principle about just distribution. But this is an incomplete view. We still need to find (X), the principle that implies (3) and is acceptable in all kinds of case.<sup>37</sup>

37. While our view is incomplete, it will face other objections. Suppose we know that, if we have children, they will have lives that are worth living, but are not wonderful. We also know that their lives will contain some suffering. On the view suggested, having these children would make the outcome worse. If we believe that causing to exist can benefit, we might reply that there is no objection to increasing suffering if our way of doing so will benefit those who suffer. But what of the risk of having children whose lives turn out to be worse than nothing, or below the Restricted Level? We might justify the taking of this risk by the chance that our children will have lives above the Wonderful Level. But this seems a weak reply. See Sikora's arguments in *Obligations to Future Generations*, op. cit., pp. 136-45.

## IX. THE MERE ADDITION PARADOX

There is another problem. Consider these alternatives:



We wish to defend the view that B is worse than A. Let us compare A with A+. The only difference is that A+ contains an extra group, who have lives worth living, and who affect no one else. This group are worse off than the group common to both worlds. If this inequality was both known, and removable, it might involve social injustice. Let us therefore assume, for simplicity, that the two groups in A+ are not aware of each other's existence, and could not communicate. A+ might be some possible state of the world before the Atlantic Ocean had been crossed. A would be different state in which the Americas were uninhabited.

Is A+ worse than A? Note that I am not asking whether it is *better*. This we have already implicitly denied, since we have denied that extra lives, if below the Wonderful Level, have intrinsic moral value. But it seems harder to believe that A+ is *worse* than A. This implies that it would have been better if the extra group had never existed. If their lives are worth living, and they affect no one else, why is it bad that these people are alive? There is one feature of A+ which seems morally regrettable. It is true here, as it is not in A, that

some people are worse off than others through no fault of theirs. There is natural inequality, or what some would call natural injustice. But if this inequality is not perceived, and involves no social injustice, it seems hard to believe that this feature is so bad as to make A+ worse than A. It seems hard to believe, simply on the ground that the extra group are worse off than some other group unknown to them, that it would have been better if the extra group had never existed.

A+ is of course worse than A on the Impersonal Average Principle. But this is one of the many cases where that Principle seems implausible. What is wrong with a lower average, if it involves no loss of any kind, but only mere addition? (The National Gallery has, in its basement, a Reserve Collection. This lowers the average quality of the paintings in the Gallery. Does this make it a worse gallery?)<sup>38</sup>

Suppose we decide that we cannot honestly claim to believe that A+ is worse than A. We must then compare A+ with B. It may help to make this comparison through the intermediate world, Divided B, where the two halves of B's population cannot communicate. Clearly Divided B is as good as B. We can now ask, if A+ were to change to Divided B, would this be a change for better or worse? On our ordinary moral assumptions, it would be a change for the better. Suppose that this change takes place within a short time. What then happens, in the move from A to Divided B, is that the worse off half gain more than the better off half lose. This is a change for the better on both utilitarian and egalitarian grounds. Remember that the worse off group are worse off through no fault of theirs. And we can suppose that they are worse off, not just because they are less happy, or have a lower quality of life, but because they have a smaller share of resources. We can suppose that B would be better than A+ both in terms of equality of welfare and in terms of equality of resources.<sup>39</sup>

38. Other objections to the Average Principle are advanced by McMahan in his "Problems of Population Theory," *op. cit.* It may be suggested that A+ is worse than A according to Rawls's *Theory of Justice*, since the worse off group are at a lower level. But where the alternative for the worse off group is not to exist at all, Rawls's theory ought, I believe, not to apply. It should not claim to handle Different Number Choices. (Rawls himself does not make such a claim.)

39. See Ronald Dworkin, "What is Equality? Part 1: Equality of Welfare" and "What is Equality? Part 2: Equality of Resources" in *Philosophy & Public Affairs* 10, no. 3 (Summer 1981) and 10, no. 4 (Fall 1981), respectively.

(Throughout this discussion quality of life, level of happiness, and share of resources are assumed to correlate, and to change together.)

It may perhaps be said that egalitarian principles only apply within some society, where there can be social injustice. But this does not seem plausible. Suppose that I know about two people—call them *Rich* and *Poor*. These two people live in different societies, which cannot communicate. Rich is better off than Poor, but is threatened by a decline in his quality of life, and share of resources. I can either intervene to keep Rich where he is, or instead help Poor. If I help Poor, I can raise him to the level to which Rich, without my help, would fall. And Poor would rise more than Rich would fall. Most of us would believe that, if I help either, I ought to help Poor rather than Rich. And we would mostly believe that this would make the outcome better. If Rich loses, but Poor gains more, this is a change for the better. This seems so even though Rich and Poor live in two societies which cannot communicate. But then Divided B would be better than A+. Therefore so would B.

Suppose we believed that A+ is not worse than A. We now believe that B is better than A+. These beliefs together imply that B is not worse than A. B cannot be worse than A if it is better than something—A+—which is not worse than A. But our first belief was that B is worse than A. So we have three inconsistent beliefs. Call this *the Mere Addition Paradox*.

This is not just a conflict between different moral principles. We may have a pluralist morality in which we believe, say, that it would be better both if there was greater equality and if there was a greater sum of benefits. There may then be cases where greater equality would lower the sum of benefits. Our two principles would here conflict. But there would be no inconsistency. We would merely have to ask whether, given the details of the case, the gain in equality is more or less important than the loss of benefits. We would here be trying to reach an all-things-considered view. In the Mere Addition Paradox, things are different. We are here inclined to believe, all things considered, that B is worse than A, though B is better than A+, which is not worse than A. These three judgments cannot all be consistently believed. They imply contradictions. One of these beliefs must go.

Which should go? Can we honestly claim to believe, of the extra group in A+, that it would have been better if they had never existed?



Or can we honestly claim to believe that a change from  $A+$  to  $B$  would not be a change for the better? If we claim the latter, we would be saying that what matters most is the quality of life of the best off people. If their quality of life falls, this cannot be made up for even by a greater gain in the quality of life of an equally large worse off group. This is so even though the worse off group are not worse off through any fault of theirs. This elitist view can apply directly to our actual world. Few of us would find it here morally acceptable.

It may be thought that, if we make the impersonal claim that only wonderful lives have intrinsic value, we have already accepted this elitism. But this is not so. We discussed that claim when considering alternatives in which different people would exist, and in each of which no one is worse off than anyone else. Suppose that lives in  $A$  are above the Wonderful Level. We might then claim that the existence of  $A$  has intrinsic value, while there is no such value in the existence of  $Z$ . This would be why  $Z$  is worse than  $A$ . When we compare  $B$  with  $A+$ , we can be considering alternatives in both of which all of the same people would exist. We might still claim here that only the lives of the best off group in  $A+$  have intrinsic value. Even though we claim this, we might believe, both on egalitarian and on person-affecting grounds, that  $A+$  should be changed into  $B$ . Our moral view would then combine principles about intrinsic value with the Narrow Person-affecting Principle and a principle about just distribution. It is a different and more elitist view that, in a world where a worse off group exists, what matters most is the quality of life of the better off group. This conflicts directly with our beliefs about justice, or equality. These beliefs do not conflict with the claims that  $Z$  and  $B$  are worse than  $A$ .

Suppose that, when comparing  $B$  with  $A+$ , we cannot accept this more elitist view. We believe that  $B$  is better than  $A+$ . If we cannot believe that  $A+$  is worse than  $A$ , we are then forced to conclude that  $B$  is *not* worse than  $A$ . If these were two possible futures for some society, it would not be worse if what comes about is  $B$ : twice as many people who are *all* worse off.<sup>40</sup>

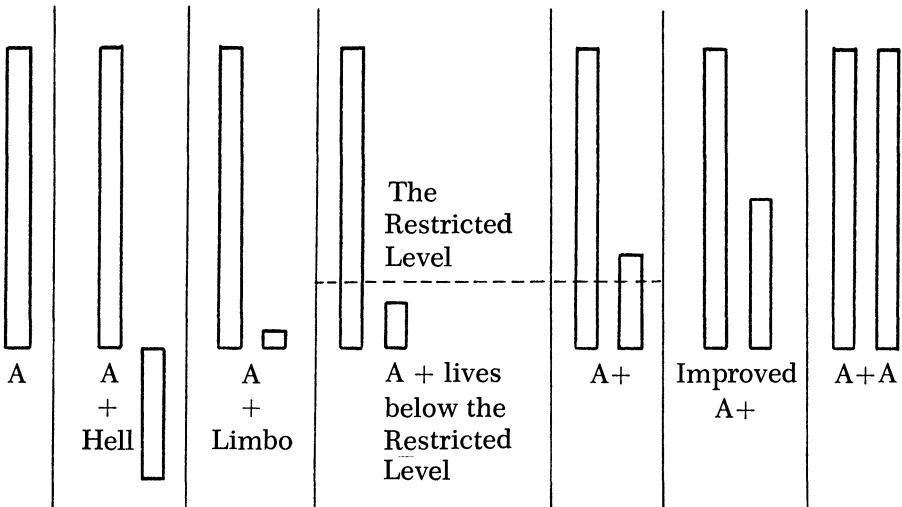
40. The argument above compares alternatives in which the same particular people would exist. We may need a further argument to carry the conclusion through to cases where, in the different alternatives, none of the same people would exist. I believe that such an argument could easily be found.

The Mere Addition Paradox does not, by itself, force this conclusion. We can avoid the conclusion if we reject one of the other two beliefs. Perhaps, though we find these hard to reject, we find it even harder to accept that B is not worse than A. Suppose we decide that, of the three alternatives, what is least hard to believe is that A+ is worse than A. We can then keep our view that B is worse than A. But we should note that we cannot simply claim that A+ *must* be worse than A, since it is worse than something—B—which is worse than A. We would here be rejecting one of the three inconsistent claims simply on the ground that it is not consistent with the other two. But this could be said against *each* claim. To solve the Paradox, we must claim that, even when we compare only A and A+, and put aside the rest of the argument, we do believe that A+ is worse. We do believe that it was bad that the extra group ever lived. To the extent that we find this hard to believe, we still have a Paradox.

It may be objected: “Your argument involves a kind of trick. When you compare A and A+, you claim that the extra group’s existence will affect no one for the worse. But by the time we have moved to B the original group have become worse off. So the addition of the extra group *is* worse for the original group. That is why A+ is worse than A.” This is a tempting reply. But it can be met. One way to show this is to restate the argument. Suppose that we are considering possible states of the world many centuries ago. There is no ground for fear about future consequences. Suppose next that A+ was the actual state in some past century. We can then ask, would it have been better if the actual state had then been A? In asking this, we can suppose that A+ did *not* later change into B. The existence of the worse off group in A+ did *not* affect the better off group for the worse. (And, since the groups could not communicate, there was no social injustice.) Given this, was A+ worse than A would have been? Was it bad that the worse off group ever lived? If we answer no, we can then ask another counterfactual question. The world did not in fact change from A+ to B. But, if it had, would this have been a change for the better? We may again seem forced to admit that it would. On this version of the argument, the ‘tempting reply’ has been undercut. The existence of the worse off group is not worse for the better off group. Since this is so, we admit that A+ was not a worse state of

affairs than A would have been. And we admit that a change from A+ to B would have been an improvement. From these two claims it follows that B would not have been worse than A.<sup>41</sup>

It may be objected next that A and A+ are not comparable. It may be said that, when we consider such imagined past states of the world, we have no view about which would be either better or worse, in some sense that has moral relevance. But consider a wider range of outcomes:



In A + Hell the extra group have (sinless) lives which are worse than nothing. If they could kill themselves, they would. Clearly, A + Hell is worse than A, and in a morally relevant way. It is also clear that A + A is *not* worse than A. Somewhere in between these two we must change our view. Where should the change come? Some would say at the level where the extra people's lives become worth living. On this view, A + Limbo is not worse than A. But we can now remember one of Kavka's principles. Kavka suggests that, if a life

41. Frances Myrna Kamm and John Mackie both suggest the view that, while it could be our duty on egalitarian grounds to change A+ into B, this change would not be an improvement. We may have a duty to do what would make the outcome worse. This view might provide a partial solution to the Mere Addition Paradox. But it would not be a full solution.

is severely *restricted*, even though worth living, it is intrinsically bad that such a life be lived. If we find this plausible, we could introduce another level, above that where life becomes worth living. Call this *the Restricted Level*. We would believe, on Kavka's view, that it is intrinsically bad if any life is lived at or below this level. Of such a life, even if it is worth living, or is of value for the person whose life it is, we would believe that it would have been intrinsically morally better if this life had never been lived.

If we accept Kavka's view, we would now have three levels: that where life ceases to be worth living, the Restricted Level, and the Wonderful Level. (None of these would be precise levels. There would be fuzziness.) And we would now have revised our partial adherence to the Impersonal Total Principle. We would believe that it is intrinsically worse if more lives are lived below the Restricted Level, and intrinsically better if more lives are lived above the Wonderful Level, but that lives between these levels have no intrinsic value.

Kavka's view provides a partial answer to the Mere Addition Paradox. We avoid the Paradox in those applications of the argument where the extra group in A+ have lives that are not above the Restricted Level. But, though this would achieve something, it would not achieve much. As I suggested, it seems hard to accept Kavka's view in the case of those whose lives are *well* worth living. The exception was the case of slaves, where there was social injustice. But this is not involved here. If we are not considering slaves, the *restricted* lives cannot be well worth living. They must be gravely deficient in all of the features that can make a life well worth living. Though worth living, they must be cramped and mean.

We can now describe another outcome, Higher Z. This is some enormous population whose lives are not much above this new Restricted Level. We would resist the claim that Higher Z would be better than A.

Can our earlier argument yield this conclusion? It may seem that it can. It may seem that, if B is better than A+, which is not worse than A, B must be better than A. By the same reasoning, C must be better than B, D better than C, and so on. But this reasoning assumes that *not worse than* here implies *at least as good as*. This is a natural assumption. But, on reflection, it seems here unjustified.<sup>42</sup> In my last

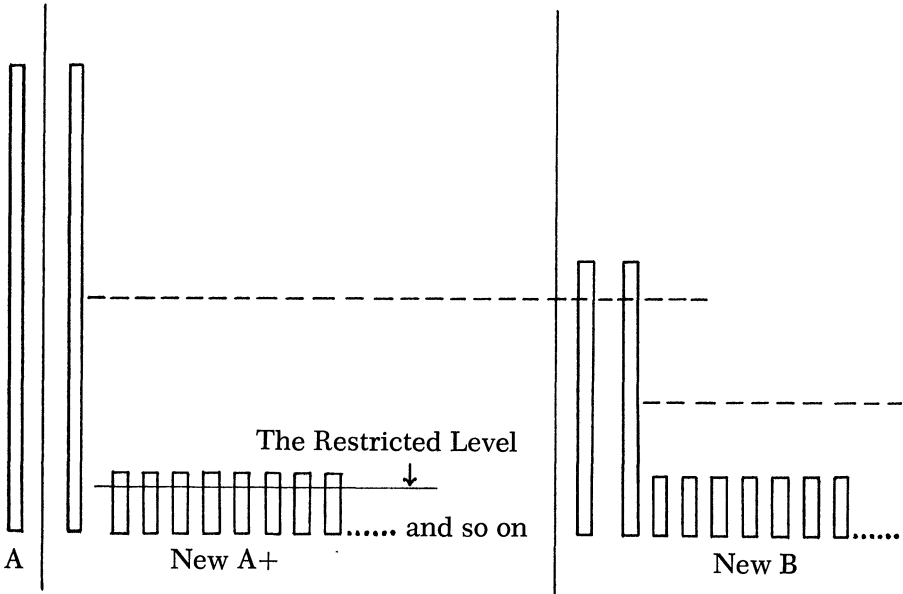
42. I owe this point to Ronald Dworkin and Amartya Sen.

diagram, compare  $A+$  with Improved  $A+$ . Clearly the latter is better (if only because it is better for the worse off group). Must we then believe that Improved  $A+$  is better than  $A$ ? We need not have this belief. We can claim that, while Improved  $A+$  is better than  $A+$ , both of these are merely *not worse* than  $A$ .

This is coherent. And in many other areas these are the kinds of claim that we ought to make. Consider three candidates for a Research Fellowship, one Historian and two Logicians. We might believe, of the Historian and the First Logician, that neither is worse than the other. This would not be claiming that they cannot be compared. It would be asserting partial comparability. There are many logicians who would be worse candidates than the Historian, and many historians who would be worse candidates than the First Logician. We are claiming, of these two, that they are in the same league. Something important can be said about their respective merits: neither is worse than the other. Next consider the Second Logician, whom we judge to be slightly better than the First. (In the case of two logicians, differences in ability can be finer.) Does this judgment force us to conclude either that the Second Logician is better than the Historian, or that the First is worse? It does not. We could claim that, though the Second Logician is better than the first, neither is worse than the Historian, who is worse than neither. Similarly, in my imagined worlds, Improved  $A+$  is better than  $A+$ . but neither is worse than  $A$ , nor is  $A$  worse than either.<sup>43</sup>

We should next note that, so interpreted, *not worse than* may not be a transitive relation. The First Logician is not worse than the Historian, who is not worse than the Second Logician. This does not

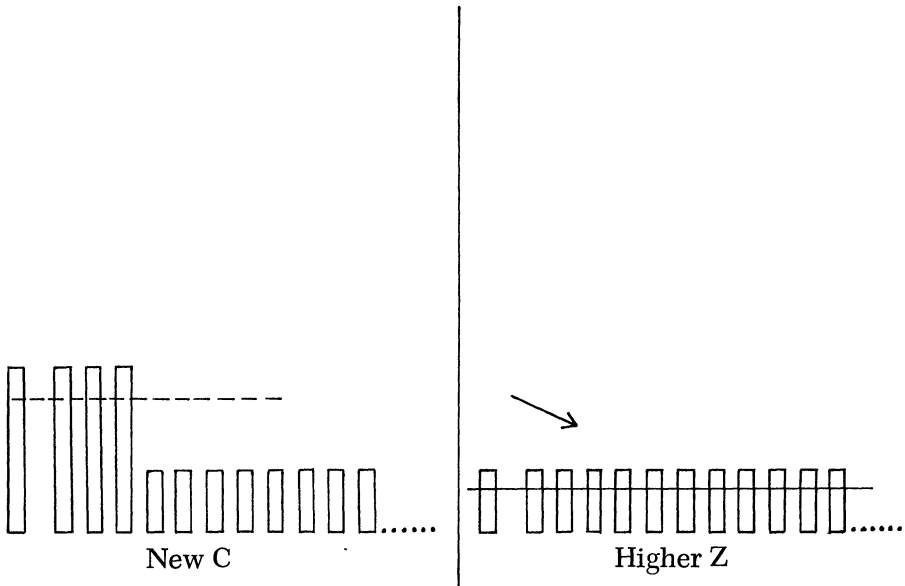
43. It may be said that, in the case of the candidates for the Research Fellowship, there is in principle full or precise comparability. Of any two candidates, the only possibilities are that one is better, or that both are exactly equally as good. On this view, the appearance of partial comparability merely results from our ignorance. If this view is correct, the case of the candidates is not the analogy that I want. But this view seems to me implausible. Must it be true, of Plato and Aristotle, either that one was the greater philosopher, or that both were exactly equally as great? This seems absurd. But it is surely true that some philosophers are greater than others, and by more or less. Such partial comparability seems to hold both for the goodness of different outcomes, and for the question of whether one person is, in morally significant ways, better off than another. See Amartya Sen's *Collective Choice and Social Welfare* (San Francisco: Holden-Day, 1967), pp. 99-102.



force us to retract our view that the First Logician *is* worse than the Second.

Because *not worse than* need not here imply *at least as good as*, we cannot move to the stronger conclusion, in the Mere Addition Paradox, that B is *better* than A. We can only conclude that B is not worse than A. And we should check that, in reaching this conclusion, we have not assumed the transitivity of *not worse than*. We have not. Our conclusion is that B is not worse than A. Suppose that it was. Suppose that A is better than B. We believe that B is better than A+. If A is better than something—B—which is better than A+, A must be better than A+. (Unlike *not worse than*, *better than* is transitive.) If we cannot honestly believe that A is better than A+, then, if we judge B to be better than A+, we cannot judge B to be worse than A.

We can reach this conclusion. But we cannot move from here even to a weakened form of the Repugnant Conclusion. It is true that, by the same reasoning, C would not be worse than B, D would not be worse than C, and so on. But since *not worse than* is not transitive, we could claim that, while C is not worse than B, which is not worse than A, C itself *is* worse than A.



But consider the alternatives presented above.

New A+ differs from old A+ by involving the existence not just of one but of many extra groups. We may imagine these as unable to communicate, perhaps as scattered on different islands. All of the people in these extra groups have lives that are not much above our Restricted Level. Their lives are such that we cannot honestly claim to believe that it would have been intrinsically better if they had never existed. This is how we defined our Restricted Level. We therefore cannot believe that New A+ is worse than A+.<sup>44</sup>

We must now compare New A+ with B. This is like the comparison between old A+ and B, except for the additional groups who are unaffected. Because there are these additional groups, it is not true here

44. It may be objected that Kavka's principle about restricted lives cannot be straightforwardly applied to my variants of A+. Perhaps we are here influenced, to some extent, by the natural inequality. Perhaps we can believe, of some extra group whose lives are *well* worth living, that their existence is morally undesirable—that it would have been better if they had never lived—*simply because unknown to them another group are better off*. I find this implausible. If we do not have such beliefs, we shall make our judgments about my New A+, and the Restricted Level, solely by reference to the intrinsic quality of the lives lived.

that, if New A+ changed to B, this would abolish the natural inequality. But it would again be true that, though the better off group loses, an equally large and worse off group would gain more. If a greater gain to a worse off group counts for more than a lesser loss to those who are better off, we must judge that New B would have been better than New A+. <sup>45</sup> Now compare New B with New C. Once again, a better off group loses, but an equally large worse off group gains more. By the same reasoning, New C would have been better than New B. In this way we could reach Higher Z. This is some enormous population whose lives are not much above the Restricted Level. Higher Z must be better than New A+, since every step down the New Alphabet has been judged a change for the better, and *better than* is transitive. We can remember next that New A+ is not worse than A. Taken together, these claims imply

*The New Repugnant Conclusion:* For any possible and large population, say of eight billion, all with a very high quality of life, there must be some much larger imaginable population whose existence, other things equal, would not be worse, even though its members have lives that are not much above the Restricted Level.

This is in two respects less repugnant than the earlier conclusion. In the original Z people's lives were barely worth living. And Z was claimed to be *better* than A. On the New Conclusion, Higher Z is only claimed to be not worse than A. But this still seems, to me at least, pretty repugnant. Lives that are not much above our Restricted Level cannot be well worth living. Even if worth living, they must be deprived of most of what gives life personal value—value to the person whose life it is. If we cannot avoid this New Conclusion, this undermines what most of us believe when we consider overpopulation. We believe that, if there were eight billion people all with a very high quality of life, this would be a better future than if there were many more people whose lives, though worth living, are gravely deprived,

45. It is often unclear whether some change makes an inequality better or worse. The question is well discussed in Amartya Sen's *Economic Aspects of Inequality* (Oxford: Oxford University Press, 1973), and Larry Temkin's Ph.D. dissertation, Princeton University, 1982. Do we reduce the inequality if we move from New A+ to New B? On most of the plausible competing views, we do.



crimped and mean—not much above the level where it would be intrinsically bad that these people are alive.

How can we resist this new argument? We must either claim that New A+ is worse than A, or deny that New B is better than New A+. The lives of those in the extra groups are above our Restricted Level. We do not think it intrinsically bad that such lives are lived. And the existence of these people affects no one else. So *why* would it have been better if these extra groups had never existed? Why is it bad that these people ever live? True, they are worse off than some other group, of whose existence they know nothing. They are, in this respect, the victims of natural injustice. But is this enough to justify the claim that it would have been better if they had never lived? Nor can we appeal to the claim that, compared with A, New A+ may have worse consequences—that, if the groups in New A+ become able to communicate, there will be social injustice, which will need to be removed, and that the result of redistribution would be worse than A. We have already seen that this reply can be blocked. We may fairly suppose that these are possible past states of the world, and that, if New A+ was the actual state, all of the others remained merely possible. New A+ did not in fact lead either to New B or to Higher Z. If this is so, we must make an intrinsic judgment about New A+ and A. Would it have been intrinsically better if the people in the extra groups had never lived? This seems hard to believe. And it seems hard to deny that New B would have been better than New A+. If we do deny this, we may have to change our view about many ordinary cases. We shall be claiming that, in a world where some people are worse off than others, through no fault of theirs, what matters most is the quality of life of those who are better off. A loss to the better off cannot be made up for by a greater gain to the worse off. It seems hard to believe this. But it seems also hard to believe that Higher Z is not worse than A. Our problem remains.

#### X. CONCLUSION: EARLY DAYS

Perhaps we can solve this problem. We could then revive the Interim Conclusions of Part VIII. We would still need to search for my desired but unknown (X), the principle that is acceptable in all cases.

There are many other kinds of case, and many other questions, that we would need to consider. I shall mention three examples.

One question is whether we could solve our problems by an appeal to human rights. I think this unlikely. But theories about rights are so many, and so varied, that this would need a long discussion.

Another set of questions concerns the quality of life. Compared with my imagined world A, world B has twice as many people, who are all worse off. But suppose that the details are these. The episodes that make a life most worth living—such as periods of good work, discovery, awareness of beauty, friendships, love-affairs—these are all, in B, of *higher* quality. People in B are worse off than people in A only because each life in B contains many *fewer* of such episodes. Their higher quality does not, for these people, sufficiently compensate for their smaller number. If these are the details of the case, we might change our view. In a move from A to B, there would not be any loss in quality. The best things in life are better in B. If we are here concerned with quality, we may thus judge B to be better than A. What would be lost in B is not quality, but quantity per person. These are very different. Many similar distinctions need to be discussed.

We must also consider the end of the human race. We know that the earth may support life for more than a billion years. And we may assume that, if we can solve the nuclear threat, mankind will have an extremely long future. If we assume this, we may feel free to judge that, in any century, it would be better if there were fewer people who were better off. The rival value, quantity, would be sufficiently provided by mankind's indefinite future. If we are to face our question squarely, we must consider cases in which the different outcomes would involve all of the people who will ever live. Only so can we be sure that we are fully considering both of the competing values: the quality of life, and the number of lives lived.

It is hard to think about the end of the human race. The complete absence of anything good may be very bad, but its badness is elusive, and our intuitions faint. This is shown by one fact about the recent debate. Consider:

*The Last Generation.* Some generation all decide that they will have no children. (Their motive might be this: Scientists discover how

to give people a thousand years of youthful life, but at the cost of sterility. All existing people choose this alternative.)

Many people have been confident that such a unanimous decision to have no children, while disturbing, cannot be open to moral criticism. When considering such a case, these people remain convinced that Wrongs Require Victims—that our choice cannot be wrong if we know that it will be worse for no one. Many of these people lose this conviction when they consider a choice like Depletion. They decide that, if we greatly lower the quality of life in the further future, there *is* an objection to our choice, even though we know that it will be worse for no one. When the Case of the Last Generation did not lead these people to deny that Wrongs Require Victims, it may seem odd that they are led to deny this by considering our choice of Depletion. The bad effect here seems, in comparison, trivial.

The explanation is that Same Number Choices, though theoretically difficult, are intuitively less puzzling. It seems clear that if the same number of people will later live, whichever policy we choose, it would be worse if those who live are worse off than those who might have lived. This seems clear enough to overthrow the view that Wrongs Require Victims. Different Number Choices are, in contrast, very puzzling. Even when the stakes are much higher, as when they involve the end of the human race, our intuitions are much weaker.

I shall not attempt, now, to summarize this paper. But it may be worth repeating two claims. First, the questions that I have been discussing do not arise only for utilitarians. They arise for anyone who believes that some people can be worse off than others, in morally significant ways, and by more or less.

My main claim has been negative. If we cannot accept either the Repugnant Conclusion, or the claim that there is no objection to our choices of Depletion and the Risky Policy, we must abandon the view that Wrongs Require Victims. More generally, we must abandon the the Person-affecting Restriction. We need some new principle of beneficence, which is acceptable in all kinds of case. Though we have not yet found this principle, we know that it cannot take a person-affecting form. It will be about human well-being, and the quality of life, but it will not claim that what is morally most important is

whether our acts will affect people for good or bad, better or worse. This reopens the question about the badness of the end of the human race. And it has implications for several smaller questions not discussed here, such as the debate about abortion, contraception, and infanticide. It is common to approach these questions by appealing to a person-affecting principle. Thus the great moral difference between infanticide and contraception is claimed to be that the first, unlike the second, is worse for an actual person. But we have learnt that, in this area, such a claim cannot provide the answer. Person-affecting principles either go astray, or are quite inadequate. They either support what we cannot believe, or cannot support what we believe.<sup>46</sup>

I shall end by mentioning the view that this area is inherently paradoxical. My desired new principle, which will solve all problems, seems to some people a mirage—not something which might be discovered. While such pessimism may turn out to be correct, it cannot yet be justified. These are early days. Judged by the number of people who have made the subject part of their life's work, rather than by the length of time since it began, non-religious moral philosophy is a very young subject. We should not be surprised that much of it is still puzzling.

46. Cf. R.M. Hare's "Abortion and the Golden Rule," *Philosophy & Public Affairs* 4, no. 3 (Spring 1975).

In writing this draft of this paper I have been greatly helped by G. Kavka, J. McMahan, R.M. Dworkin, A.K. Sen, N.E. Davis, D.H. Regan, J. Broome, J. Vickers, P. Seabright, and J.L. Mackie; also by J. Narveson, R.M. Adams, T. Scanlon, T. Nagel, S. Blackburn, L. Temkin, L.W. Sumner, M. Tooley, M. McDermott, J. Glover, S. Hurley, S. Kagan, and M. Wachsberg. To the many others who helped me with earlier drafts, I offer both thanks and apologies for my failure to keep a record of their names.