she has more than two children, whatever other people do. But if all have more than two children that will be worse for each than if none do. If these people came to see that this was true, they might achieve what I call a political solution. Though each would prefer to have more children, each might also prefer that none have more children rather than that all do. A system of rewards or penalties, aimed at stopping population growth, might be democratically adopted. Even if it was imposed undemocratically, such a system might be welcomed by all these people. Another solution would be provided by reversible sterilization after the birth of one's second child. This is a better solution, since it would impose no penalties. Once again, if they understood the facts, all these people might welcome this solution.

I have described three kinds of case. And there are other possible cases. Some would be mixtures of these three; but others would be different in other ways.

130. Overpopulation

When population growth lowers the quality of life, the effects on existing people may be either good, or bad, or neither. These effects do not raise new moral questions. But other effects do raise such questions.

These questions arise most clearly when we compare the outcomes that would be produced, in the further future, by different rates of population growth. If there is faster growth, there will later be more people, who will be worse off. As before, 'worse off' can refer either to the level of happiness, or to the quality of life, or to the share per person of resources. We should assume that, in my examples, these three correlate, rising and falling together.

Let us compare the outcomes of two rates of population growth, after one **(p.385)** or two centuries. As I have explained, there would be no one who would exist in both these outcomes. Two such outcomes are shown below.

The width of each block shows the number of people living, the height shows their quality of life. By this I mean their quality of life throughout some period. In such a period, there would be some change in the population. But for simplicity, we can ignore this fact. For the same reason, we can assume that in these outcomes there is neither social nor natural inequality; no one is worse off than anyone else. This would never in fact be true. But it cannot distort our reasoning, on the questions I shall ask, if we imagine that it would be true. And this makes my questions take a clearer form.

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 does not assume that, as my diagram suggests, these judgements could in principle be precise. I believe that

there is only rough or partial comparability. What my 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 another similarly great
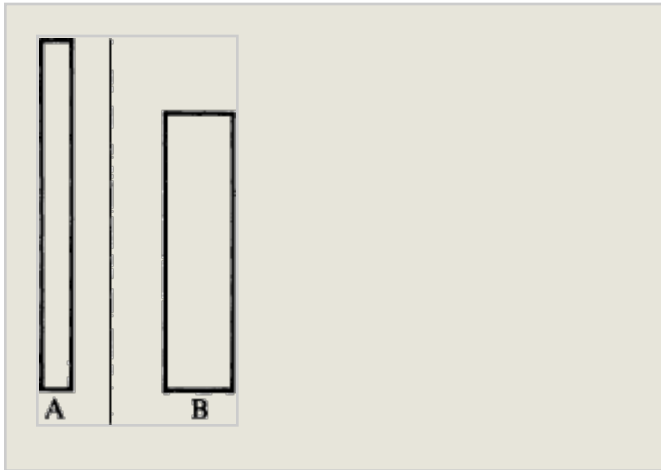
decline before people's lives ceased to be worth living.

There are various ways in which, with twice the population, the quality of life might be lower. There might be worse housing, overcrowded schools, more pollution, less natural beauty, and a somewhat lower average income. If these are the ways in which the quality of life would be lower, we can plausibly assume that it would take much more than another similar decline before life ceased to be worth living.

Except for the absence of inequality, these two outcomes could be the real alternatives for some country, or for mankind, given two rates of population growth over many years. Which would be the better outcome? By 'better' I do not mean 'morally better' in the most common use of this phrase. This applies only to persons, or to acts. But one of two outcomes can be better in another sense, that has moral relevance. It would be better, in this sense, if fewer people suffered from some crippling illness, or if the Lisbon Earthquake had not occurred. And we can clearly make such claims about **(p.386)** outcomes that involve different possible populations. Suppose that, in two such outcomes, the same number of people would exist. If, in one of these outcomes, people would be much worse off, this would clearly be the worse outcome. This outcome would be worse for no one. But, as I have argued, this does not show that this outcome cannot be worse.

Return to A and B. Which outcome would be better? It is clearly bad that, in B, people are worse off. Could this be morally outweighed by the fact that there are more people living?

Suppose we believe that, in the Case of the Happy Child, my couple have no moral reason to have this child. We may then believe that, if people are worse off, this cannot be morally outweighed by an increase in the number of people living. Those who believe this often appeal to

The Impersonal Average Principle: If other things are equal, the best outcome is the one in which people's lives go, on average, best.

Some economists make this principle true by definition.[20] I call this principle *impersonal* because it is not *person‑affecting*: it is not about what would be good or bad for those people whom our acts affect. This principle does not assume that, if people are caused to exist and have a life worth living, these people are thereby benefited.
The Hedonistic version of this principle claims

If other things are equal, the best outcome is the one in which there is the greatest average net sum of happiness, per life lived.

I state these versions in a temporally neutral form. Some state the Average Principle so that it refers only to the people who are alive after we have acted. In this form the principle implies absurdly that it would be better if, of the people now alive, all but the most ecstatic were killed. On a temporally neutral version of the Average Principle, if someone with a life worth living dies earlier, this causes people's lives to go, on average, worse.
Suppose next that we believe that, in the Case of the Happy Child, my couple do have a moral reason to have this child. We believe that it is always better in itself if an extra life is lived that is worth living. If this is what we believe, it would be natural to claim that, of my two outcomes, B might be better than A. The loss in the quality of life might be outweighed by a sufficient gain in the number of lives lived. If we make this claim, we must ask, 'What would be a *sufficient* gain?'

If we are Hedonists, we can easily state these questions more precisely. We ask

(1) 'If in one of two outcomes the people living would be less happy, can this be morally outweighed by a sufficient increase in the quantity of happiness?'

(p.387) If people are less happy, they have a lower quality of life. If we answer Yes to question (1), we must ask
(2) 'What are the relative values of quality and quantity?'

One answer is given by the Hedonistic version of the *Impersonal Total Principle*. This claims
If other things are equal, the best outcome is the one in which there would be the greatest quantity of happiness—the greatest net sum of happiness minus misery.

On this principle, B would be better than A, since in B there would be a greater quantity of happiness. Though the B‑people are each less happy than the A‑people, each of their lives contains more than half as much happiness. Since there are twice as many B‑people, they *together* have more happiness than the A‑people. (Two bottles more than

half full contain more than a bottleful.)

Suppose, next, that we are not Hedonists. What we believe to be morally important is not happiness but the quality of life. We can ask the same questions; but we must use an unfamiliar phrase. When we compare the value of quality and quantity, what is the relevant quantity? We might say, 'the quantity of lives lived that are worth living'. But this is wrong, since it ignores the *quality* of these extra lives, or how much they are worth living. The relevant quantity must, like the sum of happiness, be a function both of the number of these lives, and of their quality. To describe the relevant quantity, I suggest the phrase 'the amount of whatever makes life worth living'.

Reconsider A and B. Hedonists would claim: 'Though the B-people are each less happy than the A-people, they together have more happiness'. We can similarly claim: 'Compared with the A-people, each of the B-people has less of whatever makes life worth living. But each life in B is more than half as much worth living as each life in A. Since there are twice as many B-people, they together have more of whatever makes life worth living.'

I can now state the non-Hedonistic

> *Impersonal Total Principle*: If other things are equal, the best outcome is the one in which there would be the greatest quantity of whatever makes life worth living.
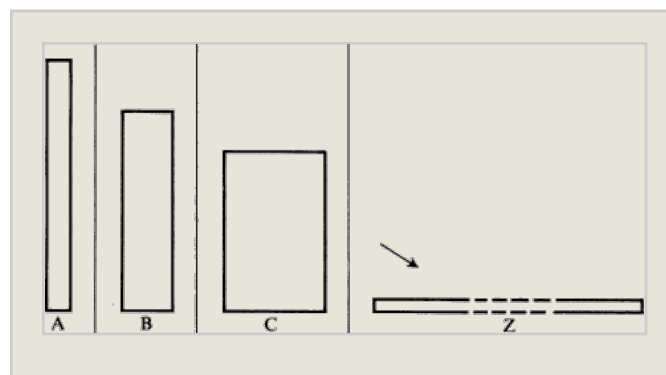
If we believe that B would be worse than A, we must reject this principle.

131. The Repugnant Conclusion

Consider next the larger diagram below. **(p.388)**

On the Impersonal Total Principle, just as B would be better than A, C would be better than B. And Z might 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 seem just worth enduring, or because it is uniformly of poor quality. Let us imagine the lives in Z to be of this second drabber kind. In each of these lives there is very little happiness. But, if the numbers are large enough, this is the

outcome with the greatest total sum of happiness. Similarly, Z could be the outcome in which there is the greatest quantity of whatever makes life worth living. (The greatest mass of milk might be found in a heap of bottles each containing only a single drop.) Let us next assume, for a reason that I shall later give, that A would have a population of ten billion. The Impersonal Total Principle then implies

> *The Repugnant Conclusion*: For any possible population of at least ten billion people, 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, even though its members have lives that are barely worth living.

As my choice of name suggests, I find this conclusion hard to accept.
A and B could in practice be real alternatives. This would not be true of A and Z. Some claim that, because of this, we need not try to avoid the Repugnant Conclusion. They might say: 'Since this conclusion does not apply to any possible choice, it can be ignored. We need not test our principles in cases that could not occur.'

I distinguished two kinds of impossibility: *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 **(p.389)** principles need to be acceptable only in the real world.[21]

It may help to remember Nozick's imagined *Utility Monsters*. These are people who get 'enormously greater gains in utility from any sacrifice of others than these others lose'.[22] 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. The world's population is now several billion. Let us imagine the wretchedness of all these people if they are denied anything above starvation rations, and all other resources go to Nozick's imagined Monster. Nozick tells us to suppose that this imagined person would be *so* happy, or have a life of *such* high quality, that this is the distribution that produces the greatest sum of happiness, or the greatest amount of whatever makes life worth living. How can this be true, given the billions left in wretchedness that could be so easily relieved by a small fraction of this Monster's vast resources? For this to be true, this Monster's quality of life must be *millions* of times as high as that of anyone we know. Can we imagine this? Think of the life of the luckiest person that you know, and ask what a life would have to be like in

order to be a million times as much worth living. The qualitative gap between such a life and ours, at its best, 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'.[23] 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 what such a life would be like, we might not find Nozick's objection persuasive. His 'Monster' seems to be a god‒like 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.

Return now 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 that 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.

We could not in practice face a choice between A and Z. Given some finite stock of resources, we could not in fact produce the greatest sum of happiness, or the greatest amount of whatever makes life worth living, by producing an enormous population whose lives are barely worth living.[24] **(p.390)** 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. Since 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 *always* be morally outweighed by a sufficient gain in the quantity either of happiness or of whatever makes life worth living. This is the question posed most clearly by comparing A 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. We have strong grounds for resisting the Impersonal Total Principle.

Someone might say: 'This is not so. This principle includes the phrase *if other things are equal*. Other things never would be equal. We can therefore ignore the Repugnant Conclusion.'

This is not plausible. What other moral principle must be infringed by the coming−about of Z? It might be claimed that this would infringe some principle about justice between generations. But this is irrelevant to our question in its purest form. We are asking whether, if Z came about, this would be better than if A came about. We could imagine a history in which only Z−like outcomes occurred. The people in Z would then be no worse off than anyone who ever lives. If we believe that Z would be worse than A, this could not here be because Z's occurrence would involve injustice.

There is another more important point. Reconsider the Non−Identity Problem. Some suggest that we can solve this problem by an appeal to people's rights. But, as the Case of Depletion shows, this is not so. If we imagine away the Non−Identity Problem, the objection to our choice of Lesser Depletion would appeal to our Principle of Beneficence. To solve the Non−Identity Problem, we must revise this principle. We must find what I call Theory X.

The same is true if we want to avoid the Repugnant Conclusion. We should not try to avoid this conclusion by appealing to principles covering some different part of morality. This conclusion is *intrinsically* repugnant. And this conclusion is implied by the Impersonal Total Principle, which is a particular version of the Principle of Beneficence. To avoid the Repugnant Conclusion, we must try to show that we should reject this version. We must try to find a better version: Theory X.

Notes:

(19) ROBERTSON, p. 460.

(19b) It is worth remarking that this case is an Intertemporal Each−We Dilemma with two special features. Since it involves different generations, the people involved cannot communicate to reach some kind of political solution, or some joint conditional agreement. And this is a Dilemma of the especially intractable kind *that includes Outsiders*.

Consider the *Auditorium Dilemma*. If the First Row stands, it will improve its view of the engrossing spectacle on stage. If it is worth standing to get this better view, it will be better for the First Row if it stands. But this would block the Second Row's view. This Row would need to stand to regain the view that it had when all were sitting. Since it would now be standing, but would not have improved its view, this outcome would be worse for the Second Row. Similar remarks apply to all the other Rows.

This case differs from an ordinary Each−We Dilemma. There are two acts: A (more altruistic), E (more egoistic). In an ordinary Dilemma, it will be better for

each if he does E, whatever others do, but if all do E that will be worse for each than if all do A. In the Auditorium Dilemma, there is a small but fateful difference. It will be better for each Row if it stands rather than sits, but if all stand rather than sit that will *not* be worse for *all* of the Rows. It will be worse for all Rows *except the First*. The First Row is *the Outsider* in this Dilemma.

Because they contain Outsiders, such Dilemmas are especially intractable. The pattern of acts that is worse for everyone else is better for the Outsiders. It would thus be worse for the Outsiders if they helped to bring about a political solution, or joined a conditional agreement. And what the Outsiders do may start a vicious chain reaction, which makes it worse for everyone to join such an agreement. Thus, in the Auditorium Dilemma, it will be worse for the First Row if all sit rather than stand. It will therefore be worse for this Row if it joins an agreement that all should sit. It may therefore stand. Once the First Row is standing, it will be worse for the Second Row if it joins an agreement that all except the First Row should sit. It may therefore stand. It will then be worse for the Third Row to join an agreement that all except the First Two Rows should sit. It may therefore stand. Similar remarks apply to every Row. The end result may be that all Rows stand rather than sit. This is worse for every Row except the First. The presence of the First Row, the Outsider, here prevents the achievement of the joint conditional agreement. And the same chain reaction may prevent the achievement of a political solution. This special feature makes such Dilemmas less likely to be solved.

Besides being trivial, the Auditorium Dilemma does not have the other depressing feature. It involves contemporaries. This makes it more likely to be solved. The other Rows might use threats to keep the First Row sitting. Or the First Row might sit merely because it expected complaints from the other Rows.

An intergenerational Dilemma does *not* involve contemporaries. This makes it harder to solve. In these Dilemmas, if all rather than none cease to give certain kinds of priority to itself, this will be better for all generations, *except the First*. The different generations cannot communicate, and reach a joint conditional agreement. Nor can earlier generations be deterred by threats from later generations. It is therefore a greater problem that this Dilemma contains Outsiders. In an intergenerational Dilemma—which need not involve population growth—the existing generation is always in the position of some Row after the earlier Rows have stood. It has already suffered from the behaviour of the earlier generations. And this earlier behaviour cannot now be altered by any political or moral solution. Since this is so, it would be worse for the existing generation if it played its part in such a solution. It would not be prompted by a reluctance to free-ride, since it cannot benefit from this solution. It would lose from its own act, and gain

nothing in return. It is thus less likely to play its part in a solution. The same reasoning will then apply to the next and all succeeding generations.

(20) See, for example, SAMUELSON, p. 551.

(21) See, for example, the discussion of the different levels of moral reasoning in HARE (1) and (2).

(22) NOZICK (2), p. 41.

(23) PLATO (2), 21 c−d.

(24) 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 the 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 casually possible net sum of utility, when they are merely used to prevent *extra* people having lives that are *worth ending* (or have net disutility).

Access brought to you by: