Problems of Population Theory*

Jefferson McMahan

I. INTRODUCTION

The eleven essays composing this volume address a wide range of interrelated questions within a domain which might be called "population theory." Among these questions are: What moral reasons might there be, given certain conditions or expectations, for or against bringing people into existence? To which principles should we appeal in determining the size, as well as the composition, of future populations? What moral reasons, if any, are there for attempting to secure the indefinite survival of the human race? Assuming that there will be future generations, what are our obligations, if any, concerning them, and what are the grounds of these obligations?

These questions are of great theoretical and practical importance. They pose extremely intractable problems for virtually all normative theories, and besides being themselves matters of urgent concern at both the individual and social levels, they are intimately bound up with a number of other widely discussed moral issues. I am persuaded, for example, that the problem of abortion cannot receive plausible treatment until certain of the questions considered in these essays have been adequately dealt with.

The essays published in this volume[1] make a valuable contribution to population theory, though their importance lies more in the exposure they give to various problems than in the solutions that they offer to these problems. Most of the authors are unduly optimistic in their belief that they have provided acceptable answers to the questions cited above,

* A review of: R. I. Sikora and Brian Barry, eds., *Obligations to Future Generations* (Philadelphia: Temple University Press, 1978), pp. xi+250; $15.00 (cloth); $10.00 (paper). I am enormously indebted to Derek Parfit for his generous help in preparing this paper, and I have profited from my acquaintance with his unpublished work in this area. I have also benefited from comments on an earlier draft by Jonathan Glover, Bernard Williams, Brian Barry, and Nick Denyer.

1. None of the essays has been previously published except for the one by Mary Warren which originally appeared in the *Canadian Journal of Philosophy* 7 (1977): 275–89.

Ethics 92 (October 1981): 96–127

though their failure in this respect is perhaps as much the fault of the subject as it is of the writers. The problems have not been solved, but it is not clear that they admit of a tidy solution.

In what follows I shall examine various of the arguments put forward in the book, and also conduct independent discussions of some of the issues that are raised. I shall first comment very briefly on some of the essays and then examine in greater detail the five which seem to me, for various reasons, to require closer scrutiny.

The best argued piece in the collection is the lengthy paper by Brian Barry on intergenerational justice. Barry's central aim is to refute one important argument for the view that obligations to future generations cannot be grounded on relations of justice between generations. The argument is based on the view that considerations of justice cannot arise where the "circumstances of justice" do not obtain. Since at least one of these circumstances, namely, the condition of relative equality, is absent in relations between generations (because future generations cannot harm us), future generations must therefore be "outside the scope of justice." This argument is assailed via what is to my mind a successful challenge to the doctrine of the circumstances of justice. Barry's paper is an attractive piece containing much good sense. I have failed to discuss it at greater length, largely because I find myself in substantial agreement with its major arguments.

The paper by Gregory Kavka also contains interesting material, though much of it is not sufficiently central to the area to merit extensive discussion in the present context. It is perhaps worth pointing out, however, that two of his major substantive claims seem not altogether harmonious. On the one hand, he contends that we ought not to add to the population if this would cause us to be unable to do as well as we otherwise might by those who already exist. He argues for this claim on the basis of an analogy with promising. If a person can fulfill only one promise, and has already made a promise, then he simply ought not to make any further promises. On the other hand, Kavka claims that there is value in the creation of extra happy people. But if he is right on this latter point, then the analogy with promising collapses, for while he thus concedes that there are moral reasons for producing new people, he appears to assume that there are no similarly weighty reasons for making new promises.

The theory advanced by Mary Warren is very similar to several of the theories, in particular Narveson's, which I shall criticize later, and is thus open to many of the objections I shall urge against these theories. Her essay is accompanied by a reply by William Anglin, the latter portion of which ably criticizes certain aspects of her theory.

The practice of discounting the utilities of future people is the subject of two of the papers, one by Robert Scott, Jr., who defends the practice, and the other by Mary Williams, who attacks it. I shall later suggest why I think Scott's argument is unsuccessful.

II. SCHWARTZ AND THE IDENTITY PROBLEM

Thomas Schwartz's piece is largely concerned with what Derek Parfit has called "the Identity Problem."[2] As this problem is crucial for the arguments of a number of the papers in this volume, it will be well to state briefly what it is. The pivotal claim is that the identity of a person is in fact (on any plausible viw of this matter) dependent on the timing of his conception. If a person's mother had conceived at a different time, she would, in fact, have had a different child. A second claim is that the implementation of any large-scale social policy must, because of its widespread though perhaps minor effects on people's lives, affect the timings of numerous conceptions. It follows that if a major social policy is put into effect, there will exist as a result numerous people who would not have existed if the policy had not been adopted. After several generations have passed, it is likely that there will be no one alive who would have existed otherwise. Suppose now that the policy in question is one that ensures that the lives which will be lived at this future date will be substantially less worth living than the lives which would have been lived had the policy not been implemented. The upshot of the Identity Problem is that the people whose lives will be of this relatively poor quality cannot complain that they have been harmed[3] or that they are worse off than they might have been had we chosen some other policy; for if any other policy had been adopted they would not exist. Thus, social policies which reduce the quality of life in the further future cannot be criticized on the ground that they are against the interests of the particular people who will live at that time.

Parfit originally formulated this argument as an objection to moralities which take what he calls a "person-affecting" form, moralities which are concerned only with the effects of acts on particular people. These moralities cannot directly criticize social policies of the sort just mentioned, and thus would appear to be unacceptable. Indeed, Richard Sikora, in his contribution to the volume, claims that any theory which has this sort of consequence "can be dismissed out of hand" (p. 126—see the whole of his sec. 5).

The position which Schwartz defends has this consequence—that is, it permits us to ignore a policy's distant effects on the quality of life. It has this feature because it insists that we cannot "be *morally required* to adopt [a social policy] unless we owe it to someone to adopt [it]—unless our *not* adopting [it] would, in some broad sense, *wrong* someone"[4] (pp. 11-12). If a future person would not have existed had we not adopted a certain policy, we cannot have owed it to him to adopt some other policy instead;

2. Derek Parfit, "Overpopulation, Part I" (unpublished). Also see his "On Doing the Best for Our Children," in *Ethics and Population*, ed. M. Bayles (Cambridge, Mass.: Schenkman Publishing Co., 1976), p. 101.

3. The claim that these people have not been harmed is open to question in cases in which their lives are worse than no life at all. I shall return to this question later.

4. These claims are not equivalent. See Sec. III, n. 9.

nor could we (with the possible exception of cases of the sort mentioned in n. 3) have wronged him by not adopting this other policy.

The fact that this position has the implication to which Parfit has called attention is not something about which Schwartz is embarrassed or apologetic. On the contrary, it is the central thesis of his paper that we have no obligation to avoid reducing the quality of life in the further future (p. 3). (He later qualifies this claim.) Those who have looked through Dennett's "Philosophical Lexicon" may recall the satirical definition of the verb "outsmart": "To embrace the conclusion of one's opponent's *reductio ad absurdum* argument." Many of us may feel that Schwartz is "outsmarting" us. We would reject his position on moral requirement before we would acquiesce in his conclusion. This choice is made easier by the fact that his position is undefended. The only reason he gives for thinking that our obligations must be owed to particular people is that he cannot see how it could be otherwise (p. 11). A review of the arguments in the succeeding papers may incline us to skepticism regarding the prospect of finding a satisfactory alternative, but it nevertheless seems clear that the inquiry into our "obligations to future generations" does not end, but only begins, with the recognition of the Identity Problem.

III. NARVESON AND THE PERSON-AFFECTING RESTRICTION

Jan Narveson's contribution to the volume under review is the most recent member of a series of seminal papers which he has devoted to the issues with which the book is concerned. Explicit commitment to the person-affecting as opposed to the "impersonal" version of Utilitarianism appears first in Narveson's earlier work. This most recent effort constitutes an attempt to modify his original theory in response to certain objections. The resulting theory, however, incorporates so many diverse strands of thought, which are not clearly distinguished and which cohere only doubtfully, that it is hard to tell whether the revised theory marks an advance over the original one or not. In this section I shall explore the connections among the various elements in the revised theory and then attempt to judge how successful the revision has been.

The theory which Narveson presents himself as defending is based on an understanding of welfare according to which consideration should be restricted to effects on people for better or worse. This will be referred to as the "person-affecting restriction." It can be a component in a pluralist, nonutilitarian theory, or, as in Narveson's case, it can form the basis of a version of Utilitarianism. Narveson originally formulated this version of Utilitarianism in order to avoid certain implications of the "Total View"[5]—the version of Utilitarianism according to which what matters is

5. The phrase is Narveson's. This version of Utilitarianism is based on a "totaling" principle which can also be a component of a nonutilitarian morality. It should be emphasized that problems of population theory are not peculiar to Utilitarianism (though they do not arise quite so acutely for some theories—e.g., Libertarianism).

the total amount of utility in the world. The Total View implies, inter alia, that it could be better to expand a population even if everyone in the resulting population will be worse off than everyone in the original, and that it will normally be wrong for a person to remain childless. Narveson's theory avoids these implications if it incorporates the claim that a person cannot have been affected for better or worse, or been benefited or harmed, by being brought into existence. I shall discuss this claim, about which Narveson is curiously ambivalent, in the next section.

This theory has encountered certain problems. Consider the view that, while the fact that a person's life would be worse than no life at all (or "worth not living") constitutes a strong moral reason for not bringing him into existence, the fact that a person's life would be worth living provides no (or only a relatively weak) moral reason for bringing him into existence. This view, which I shall refer to as "the Asymmetry," is approved both by Narveson and by common sense. But since Narveson's theory incorporates the claim that conceiving[6] a child does not affect him for better or worse, it can account only for the second component of the Asymmetry. We will see later why Narveson thinks that his theory can justify the whole of the Asymmetry. A second problem (which Narveson unfortunately runs together with the first) is that his theory cannot require a couple to do what would cause them to have a happier rather than a less happy child where, because of the Identity Problem, these children would be different. If, for example, a woman has a temporary condition which would cause any child she conceives now to be handicapped, Narveson's theory could not require her to postpone conception until her condition is corrected. Narveson thinks, however, that this second challenge can be met by simply introducing the following qualification: if one has decided to have a child, one should choose to have a happier rather than a less happy child where this is possible. Thus Narveson summarizes his revised theory as follows:

> (1) New additions to population ought not to be made at the expense of those who otherwise exist, even if there would be a net increment in total utility considered in person-independent terms. But (2) new additions ought to be made if the benefit to all, *excluding* the newcomer, would exceed the cost to all, *including* him or her, as compared with the net benefit of any alternatives which don't add to population [i.e., if the benefit minus the cost would exceed the net benefit of any alternative]. Finally, (3) within those limits, the decision whether to add to population is up to the individuals involved in its production, provided that if they have a choice of which child to produce they produce the happier one, other things being equal. [Pp. 55–56]

Principle 1 simply states a consequence of the person-affecting restriction. Principle 2 is clearly intended to incorporate the Asymmetry into Narveson's theory, though it does not, in fact, articulate the Asymme-

6. I assume that "conception" is acceptable as shorthand for "coming into existence."

try unless "if" is read as "if and only if." Principle 3 contains the addition cited above.

There is an obvious problem here. In this essay and elsewhere, Narveson is explicit in affirming the claim that conception cannot be a benefit or a harm (e.g., p. 48), yet it is clear that this claim is incompatible with 2—that is, with the Asymmetry expressed in person-affecting terms. The notion of the "cost to the newcomer" must be understood to mean the extent to which the newcomer is harmed by coming to exist. The only alternative is to construe the "cost" impersonally, as an increase in the amount of misery in the world; but this alternative involves giving up the person-affecting restriction.

How is it, then, that Narveson thinks that he can establish the Asymmetry on the basis of his theory? The reason he thinks this is, I suggest, that he tends to rely on the view that an act cannot be wrong unless there is some actual person for whom it would be worse. There must, in short, be a complainant. There is abundant evidence to suggest that Narveson accepts this requirement (e.g., pp. 43, 50, and 55–56). I shall cite only one example. On page 47 Narveson gives what he takes to be the reason why his theory does not impose an obligation to produce a happy child. One would expect him to say that the reason is that conceiving the child would not count as affecting him for the better. But in fact what he says is that "if we do not have it, it is not the case that the child in question will be worse off, because it won't *be* at all." If Narveson is here appealing to the requirement that there be a complainant, then in order to establish this component of the Asymmetry he does not even need the claim that conception cannot be a benefit or a harm. Indeed, if he were to abandon that claim altogether, he could then support the whole of the Asymmetry on the basis of this requirement.

This requirement is not only compatible with the person-affecting restriction, but can even be regarded as one way of interpreting it. In fact, if we decide that conception can be a benefit or a harm, then there are two possible ways of understanding the person-affecting restriction. These correspond to two different ways of construing the phrase "worse for people." One interpretation keeps the requirement that there be a complainant. On this interpretation, an act is worse for people only if there are or will be people for whom it is worse. On the other interpretation, an act is worse for people if some other act would have been better for people.[7] I shall call these two interpretations, respectively, the narrow and broad versions of the person-affecting restriction. These yield two corresponding versions of Utilitarianism. The difference between these should not be confused with the difference between classical and negative Utilitarianism.[8] *Both* of these versions can require us to benefit people.

7. These two claims, stated somewhat differently, are illuminatingly discussed by Parfit in "Overpopulation, Part I."

8. Narveson confesses that he feels "an uncomfortable presentiment" that he is being influenced by the view that "our duty is only not to harm people, rather than to promote

The difference is that, on the narrow version, we ought to benefit people only when, if we fail to do so, there are or will be people for whom this is worse. These are the "complainants." On the broad version, no complainant is required. To see this, suppose that the act which would be best for people is one which brings new people into existence, and that these latter people would be the ones affected for the better. If this act is not performed, there will be no one for whom this is worse, but it will still count, on the broad version, as worse for people. I shall return to this version later.

Narveson never actually states the view I have attributed to him. But he seems to be gesturing toward it when he claims that "duties must always be duties to someone" (p. 43). This claim will not, however, do the required work. Rather than supporting the Asymmetry, it is actually incompatible with it. On this view there could not be a duty not to produce a miserable child, for there would be no one to whom it would be owed. It seems to me that what Narveson actually has in mind here (or at least what he needs) is the claim that a duty (or perhaps an obligation) must be such that, if it is violated, there is someone who is thereby wronged.[9] This claim is analogous (though not equivalent) to the narrow version and similarly supports the Asymmetry.

If Narveson were to adopt the narrow version (or the analogue just cited) he could then explain the Asymmetry and thus preserve his principles 1 and 2. He could thus avoid the first of the two objections cited earlier. But he could not avoid the second; for he would be unable to preserve 3. For if a person were to disobey 3, the narrow version would judge that his act was not worse in person-affecting terms. Narveson recognizes that if 3 is disobeyed "no existing person is [thereby made] worse off than he or she might have been" (p. 55). He seems to think, however, that this problem can be surmounted by pointing out that "one person is worse off than another might have been if the other option had been taken"; but it is hard to see, in the light of Narveson's theoretical posture, how this observation is relevant.

The alternative is to return to Narveson's original theory with its insistence that conception cannot be a harm or a benefit. This theory, however, avoids neither of the two objections. The theory cannot support the Asymmetry, and principle 3, which is a concession to the impersonal theory introduced to avoid the second objection, violates the person-affecting restriction.

Neither theory, therefore, avoids both objections. Moreover, both theories have the same implications for social policies, exposed by the Identity Problem, which Schwartz's position has. The choice between the two is, thus, not very tantalizing, and I can think of no third way in which

---

utility however and whenever" (p. 56). If, however, he is appealing to something like the narrow version, this suspicion is unfounded.

9. This and Narveson's own claim correspond to Schwartz's two claims about moral requirement. That Schwartz seems to regard the two as equivalent suggests that Narveson may be doing the same.

Narveson can secure the coherence of the elements he wants his theory to encompass. Since he is attempting to defend commonsense intuitions, this is an unwelcome result.

Let us suppose, however, that 1, 2, and 3 can all draw support from some unknown source. Is the theory comprising these three claims acceptable?

Narveson's theory is designed to deal with two types of case. Of the first sort are cases in which we must decide whether to add to the existing population or not. These cases are governed by principles 1 and 2. Cases of the second sort, which are governed by principle 3, are those in which it has been determined that some addition to the population is to be made, though it must be decided which particular addition this will be. Narveson confines his attention to cases involving the possible additon of a single person. But the range of cases which Narveson's theory must cover is not as narrowly circumscribed as he seems to assume. While it is true that, at least at present, only parents can produce new people, there is clearly a sense in which members of a community concerned with population policy may also be said to face decisions concerning the possibility of adding to the population. Thus, we need to consider not only the possibility of adding a single individual, but also the possibility of adding an entire group. At this point a problem arises.

Suppose that population planners have, in accordance with Narveson's principles 1 and 2, determined that it will be advantageous to adopt a policy which would encourage people to bring a large number of new people into existence. And suppose further that there is a choice between two policies, each of which would result in the production of a different group of people. While one group would be more numerous and would, thus, have a greater total of happiness, the other would be less numerous and would have a higher average level of happiness (because, for example, each member of the smaller group might be better off than anyone in the larger). The two groups would, moreover, have no members in common. Finally, let us assume that there would be no advantage to existing people in the production of one group rather than the other. Principle 3 would presumably instruct the planners to opt for the happier of the two groups; but which one is this?

At first it would appear that Narveson would favor the group with the higher average, since his "concern is that whatever people there are be as happy as possible. . . . The concern that there be more people, simply to maximize instances of happiness in the universe, seems," he writes, "of another order" (p. 55). But he is, in fact, debarred from appealing to an averaging rule, since he accepts the objection which Parfit has urged against the Average View (the version of Utilitarianism which instructs us to maximize average utility), and this objection is equally applicable in the present case.[10] Thus, Narveson is left without a principle for evaluating outcomes in which different numbers of people exist, and his theory is,

10. See Narveson's "Moral Problems of Population," in Bayles, p. 74. Parfit's objection is briefly stated in Sec. VIII of the present paper.

as it stands, incomplete. For most moral theories, incompleteness on this point may not be regarded as a very grave defect, since the problem is not one that men have often faced, or one about which we have very decided intuitions; but it does seem a lamentable shortcoming in a theory that purports to deal with population problems. Of course, Narveson might simply say that, as long as effects on existing people are equal, it does not matter which group is produced (except perhaps when it is obvious that one group is "happier"—for example, when it is preferred on all common criteria: total, average, maximin, maximax, etc.). I doubt, however, that anyone who finds 3 an important component of Narveson's theory will be satisfied with this response.

IV. THE PROBLEM OF CONCEPTION

Thus far I have argued not only that there is an important objection to the theory comprised by Narveson's three principles, but also that this conjunction of principles cannot be supported by either Narveson's original theory or the narrow version of the person-affecting restriction. Both of these latter theories, moreover, are themselves open to serious objections. If, however, Narveson were still tempted to choose between them, one consideration which would certainly be relevant to his choice is the question whether the claim that conception cannot be a benefit or harm is true. The claim is, of course, of interest and importance apart from its relation to Narveson's theoretical concerns, and for this reason it is regrettable that it is hardly touched on in the book. The only writer who mentions it is Narveson himself, and his discussion of it is relatively cursory. I propose, therefore, to attempt to fill the lacuna by briefly examining this claim. I shall approach the topic by way of Narveson's comments.

Although, as we have seen, Narveson's argument at one point presupposes the rejection of the claim, he nevertheless professes to embrace it. He contends that, "since conception is not something which happens to someone who is already around, . . . we cannot say that the resulting person is better off than he or she would have been if that event had not taken place" (p. 48). This claim presupposes what Narveson calls "the comparative view," that is, the view that effects on a person's welfare can be evaluated only by comparison with alternative states in which the person would be better or worse off. But this view has the consequence that the effects of saving lives cannot be evaluated, since death is not a state in which a person is worse off. To avoid this implication, the comparative view should be restated so that the relevant comparison is with alternatives which, for the person concerned, are better or worse, or more or less desirable. Narveson's claim could then be that, while death is normally worse for a person, or less desirable for him, than continued existence, there is no one for whom nonconception is worse. Thus, on the comparative view, one can affect people for the better by saving them, but not by conceiving them.

There are two ways of resisting this argument. One is to contend that,

from the perspective of an actual person, nonconception can be viewed as an alternative which would have been worse (or better), not just impersonally, but for that person. The fact that people do manifest attitudes of gladness or regret concerning their own conceptions suggests that this contention is at least intelligible.[11] The contention should, moreover, be compatible with the obvious truth that, in cases in which conception does not in fact occur, this is not worse for the person who might have existed. The difference between these two contentions is that while the latter implies that there is never any subject to whom any misfortune could be ascribed, the former presupposes that the subject is entrenched in the terra firma of actuality.

It seems possible, on the basis of the foregoing contention, to reason as follows: "If we produce a new person, he can then claim that his nonconception would have been worse for him, and therefore that, on the comparative view, his conception was better for him. Thus even on the comparative view we would have benefited this person by bringing about his existence." On the other hand, however, it seems equally admissible to reason on the basis of the opposite assumption: "If we do not produce the new person, this will not be worse for him, since he will then never exist. Thus, on the comparative view, it cannot be better for him to be brought into existence." In short, if we accept the comparative view, we seem to be faced with the following paradox: if not producing a new person would not be worse for him, then producing him would not be better for him—though if he is produced, then it can be said that not producing him would have been worse for him, so that producing him was better for him.

The preceding argument seems both obscure and inconclusive. The second response to Narveson's argument is more straightforward: it is simply to reject the comparative view. The reason that this view tends to command ready assent is that the case which presents the most radical challenge to it—namely, conception—has, until recently, rarely come up for discussion. Thus, there has been little occasion for questioning its validity. But suppose that in the case of conception the requisite comparisons cannot be made. Then, if we think that there is value in the creation of happy people, or negative value in the creation of miserable people, and if we are dissatisfied with the impersonal account which holds that these acts (or rather their consequences) have value only "from the point of view of the universe" rather than that of the people created, we may be inclined to dispute the necessity of comparison. While we might concede that phrases such as "better off" and "better for" are essentially comparative, and thus that conception cannot affect a person for better or worse, nevertheless it would still be possible to contend that conception can affect a person "noncomparatively"—that is, it can affect him for good or ill even though the alternative would not have been worse or better for him.

11. "Let the day perish wherein I was born/ And the night when it was said/ There is a man-child conceived" (Job 3:3).

Narveson considers the possibility of abandoning the comparative view, but appears to find this move unsatisfactory:

> The best one could do is set the utility-state of [a] possible person at zero if not conceived, and then say that the person's conception has benefited him to, say, the extent of the contribution of his life-long utility attributable to his genetic make-up as opposed to other contributing factors. But this is to equate the condition of someone who is neither happy nor unhappy with that of one who is dead, and that seems wrong; worse yet, it is to equate it with "one" who was never born in the first place, and thus has no identity at all. And that seems very strange indeed. [P. 48]

Narveson's central objection to dropping the comparative view is, thus, connected with the problem of assigning a value to nonexistence—in particular, the problem of ranking nonexistence in relation to a scale of utility-value. His uneasiness might, however, be to some extent dispelled if he were to dispense with certain assumptions which he appears to make. For one thing, his language is suggestive of a hedonistic understanding of utility, so that the implication of the passage just quoted might be that, if nonexistence is assigned a value of zero, then nonexistence can be no better or worse than a state of hedonic neutrality (cf. Sec. VII). Since Narveson is avowedly not a hedonist, he might regard this observation as a *reductio* against hedonism rather than as a point against rating nonexistence at zero, and in this I think he would be right. It seems to me that a life with a utility-value of zero should be understood as a life which, for whatever reasons,[12] is neither preferable to nor any worse than no life at all. On this view it seems reasonable to assign a value of zero to nonexistence.

Another factor which seems to contribute to Narveson's reluctance to embrace the view that nonexistence has a value of zero is the assumption that this view requires the assignment of "utility-states" to people who do not exist. This ground for hesitation can also be removed. Nonexistence involves the absence of utility, positive or negative, and if utility is what has value (or what constitutes value), then nonexistence has no value. Thus, in aligning nonexistence at the zero point on the utility-scale, we are thereby enabled to begin calculating on a blank slate when we attempt to estimate the extent to which a person has been benefited or harmed by being conceived (or, alternatively, the extent to which his conception will increase or decrease total utility—excluding effects on other people). Calculating on a blank tablet does not, however, involve attributing a utility-state to a possible person. It simply registers the fact that prior to the person's conception there was, so to speak, a utility-vacuum which that person's life has subsequently filled.

Matters are perhaps complicated, however, if we accept an analysis of utility in terms of the satisfaction of desire, for then it might be urged that

12. This understanding is thus compatible with, but not favorable to, a hedonistic analysis of utility.

there are numerous instances in which nonexistence does have a certain utility-value—namely, those cases in which a person's nonexistence constitutes the state of affairs the obtaining of which satisfies or frustrates one of his desires—for example, the desire to die or the desire not to die. This complication arises, however, only in connection with posthumous nonexistence, since neither the total nor the prenatal nonexistence of a person can, when it obtains, constitute the state of the world which satisfies or frustrates a desire of his. Thus, the preference theory of utility poses no obstacle to setting the value of prenatal nonexistence at zero, and thus cannot be invoked to reinforce Narveson's objection to the abandonment of the comparative view.

Thus far I have argued against Narveson's defense of the claim that conception cannot be a benefit or harm. I do not pretend, however, to have shown that this claim is false, but only that it is not obviously true. Nevertheless, if we suppose that it is false, a further problem arises: how should we determine the amount by which a person is benefited or harmed by being conceived? The assumption which is usually made by adherents of the Total View is that the act of creating a new person should be credited with increasing or decreasing the total utility by the net amount of positive or negative utility which the new person's life can be expected to contain (not counting effects on others). Should we make the parallel assumption that conception benefits or harms a person to the extent of his lifelong balance of utility, or should we assume that the effect is of a lesser magnitude?

Narveson's view, as we have seen, is that conception could be said to benefit a person only by the amount of his utility attributable to his genetic endowment: " . . . Conception, as such, contributes only a small fraction of the utility of the resulting person, most of which is due to ongoing activities of parents, etc., etc." (pp. 53–54). He does not argue for this view, but simply registers his agreement with L. W. Sumner, whose argument is—presumably, for no reference is given—contained in an unpublished paper, "Morality and the Production of Life" (pp. 24–28). Sumner argues against the view that the benefit (or harm) which conception confers is equal in amount to the net total of utility in the person's subsequent life on the basis of an analogy with voting. He notes that, just as conception "is a necessary condition of all the happiness that [a person] will enjoy," so in a case where 100 votes are required to win an election and exactly 100 votes are cast, each vote is a necessary condition for winning. But, just as it would (allegedly) be absurd to credit each vote with the whole of the benefit derived from winning the election, so it would likewise be absurd to credit conception with the whole of the benefit for which it is necessary (but not sufficient). To determine how credits should be assigned, Sumner recommends the general "method of factoring the causal contributions of those agents whose actions were jointly sufficient to bring about the result" (p. 26). Thus, in the case cited, each voter presumably contributes one one-hundredth of the result.

What is puzzling about this view, however, is that it leads one to expect that, if one were to subtract the contribution of one of the voters, one would be left with 99 percent of the original outcome. But a person cannot be 99 percent elected. Analogously, it suggests that the total amount of benefit in a person's life would have been only slightly less in the absence of the fractional contribution to that sum which, according to this view, was made by his conception—just as it would have been slightly less had he not been treated to dinner the other evening. But again this apparent implication is obviously false. There is, moreover, a further, related objection to Sumner's view which is pressed by Sikora in his essay. Sikora's own example is rather exotic, however; so I will present an analogue of his argument, using the example of conception itself. Suppose that it is open to an agent either to conceive a child or to benefit an existing person. (Assume that in each case both the cost to the agent and effects on others would be equal.) While the child would subsequently be very happy, the amount of this happiness attributable to its conception would, according to Sumner's view, be only a fraction of the total—indeed, it would be less than the amount of benefit which the agent could bestow on the existing person. If the agent's concern is to maximize benefit then, according to Sumner's view, he ought to benefit the existing person, even though the enormously greater amount of benefit which the child could have enjoyed will thereby be lost.

Sumner's view fails to distinguish between cases in which a person's contribution is indispensable to the production of a desired outcome and cases in which it is not. Consider, for example, a case in which each of 100 people gives a dollar to charity. Sumner's view yields the right answer in this case. But it implies that the contribution of each donor in this case counts for just as much as the contribution of each voter in the voting case, and this is implausible since, while the lack of one dollar would not diminish the value of the others, the lack of one vote would neutralize the efficacy of the other ninety-nine.

The foregoing arguments seem to indicate that Sumner's view is deficient in cases in which some or all of the factors involved in the production of an outcome are necessary. Conception is such a case. It would seem, moreover, that these objections to Sumner's view will hold against any view that ascribes to conception only some fractional contribution to a person's subsequent utility. Nor will it do to claim that, since conception is not itself sufficient for any benefit to the person conceived,[13] it therefore contributes nothing, for the same could be said of each of the 100 votes. Thus, if only by default we seem to be left with the view that conception benefits or harms a person to the full extent of his overall balance of utility.

This view, in fact, seems plausible. Suppose that we accept the person-affecting restriction and also believe that conception can be a

13. To this extent, Narveson is right to mention "the ongoing activities of parents, etc." These are also necessary on each occasion of some particular benefit.

harm. If a person would be utterly wretched throughout the whole of his life if he is conceived, then according to the view Sumner and Narveson advocate, only a fraction of the net amount of misery he would experience should be weighed against the benefit to others of having him conceived. But according to the view I recommend, if the net amount of harm he would subsequently suffer is greater than the amount by which people will be affected for the worse if he is not conceived, then he ought not to be conceived. This seems the more plausible view to take, since it is this amount of harm, and not some fraction of it, that will be prevented if the person is not conceived.

I shall conclude this section by reviewing where the various theories discussed in connection with Narveson's essay now stand. Narveson's original theory seemed inadequate for several reasons. This judgement is reinforced if we decide that the claim that conception cannot be a benefit or harm should be rejected. The narrow version of the person-affecting restriction also faced strong objections. The other alternative mentioned was the broad version. Parfit has claimed that this version restates the Total View. And if what I have argued is right—that is, if conception benefits a person to the extent of his lifelong balance of utility—then Parfit's claim is correct: the broad version is extensionally (though not formally) equivalent to the Total View. These conclusions do not force us to accept the Total View, for there are other possibilities—for example, the Average View—but they do not augur well for the attempt to find an acceptable way of avoiding the implications of that theory within a consequentialist framework.[14]

V. BENNETT'S THEORY

A dominant concern of Jonathan Bennett's essay is to defend what, in Section III, I referred to as the Asymmetry—the claim that while the fact that a person would have a life worth not living provides a reason not to conceive him, the fact that a person would have a life worth living does not provide a reason to conceive him. His defense consists in laying down the following rule: "The question of whether action A is morally obligatory depends only upon the utilities of people who would exist if A were not performed" (p. 62).

This rule prevents us from counting, as a reason for doing action A, the utilities of people who would exist if and only if A were performed; thus, it provides the desired claim that there can be no obligation, based on a person's probable utilities, to bring that person into existence. Now it would seem that, by parity of reasoning, we should also be prevented from counting, as a reason *against* doing A, the utilities of people who would exist if and only if A were performed. But this is not the case. Bennett's rule requires that forbidden acts be described in terms of the language of "not doing." This requirement entails that the final clause of

14. This paragraph summarizes some of the main conclusions of Parfit's "Overpopulation, Part I."

the rule (where "A" has been replaced by a description of the act) will, in these cases, always contain an instance of double negation, so that one is allowed to consider the consequences which ensue when the act in question is performed. Take, for example, the case of conceiving a miserable child. If we replace A with "not conceiving a miserable child," Bennett's rule reads: "The question of whether not conceiving a miserable child is obligatory depends only upon the utilities of people who would exist if the act of not conceiving a miserable child were not performed (i.e., if the child were conceived)." This structural feature of his rule permits Bennett to contend that, if people come into existence as a result of a certain act, their utilities can count as a reason for forbidding the act, but not as a reason for requiring that it be performed.

Bennett supplies two reasons for thinking that his rule captures the form in which claims about utility should be expressed: "The main reason . . . is that it is harmless . . . " (pp. 63 and 72). In the area of morality with which the essays in this volume are concerned, a "harmless" rule would indeed be a welcome discovery; but Bennett is deceived to think that his rule merits this description. His rule is structurally analogous to the narrow version of the person-affecting restriction, and thus inherits the objections to that theory which we have considered in connection with Schwartz's and Narveson's essays. The rule is therefore not innocuous, and the main reason for accepting it collapses.[15]

Bennett has, however, a further reason for thinking his rule correct, which is based on the contention that "no sense attaches to '(dis)utility for a person who is at no time actual.'" He contends that philosophers who rely "on the notion of amounts of possibly unowned happiness" are guilty of "a philosophical mistake": "As well as deploring the situation where a person lacks happiness, these philosophers also deplore the situation where some happiness lacks a person . . ." (pp. 61, 64, and 68). Bennett believes that his rule allows him to avoid this sort of "bad philosophy" (p. 70).

It seems obvious, however, that, rather than providing grounds for the acceptance of the rule, the exposure of this philosophical mistake (if it is one) should persuade us to reject the rule. If, in accordance with the rule, we judge that a certain person must be prevented from coming to exist on account of the negative utility which he would have to experience, are we not guilty of the alleged absurdity of taking account of the negative utilities of "a person who is at no time actual"? When we applaud a decision not to produce a miserable child, are we not welcoming a situation in which some misery lacks a person? In order to maintain his stance on "unowned utilities," Bennett must hold that we cannot say that

15. It might be thought that the rule would also encounter difficulties in judging it obligatory to refrain from killing, since it requires that the obligation be based on consideration of the utilities of those who would exist if the person were killed, thereby excluding the utilities of the victim. This objection can, however, be met by stipulating that the word "exist" must be understood tenselessly—as referring to all those who exist at some time. (This reply is similar to the reply which can be made to a common argument against the Average View; see Sec. VI.)

it would be (or would have been) wrong to produce a miserable person; we could recognize and condemn this crime only when it is actually committed. This is not only an odd result, but it is also at variance with the deliverances of his rule.

The "philosophical mistake" is, therefore, presumably not a mistake. But if Bennett's rule is rejected (on account of the implications exposed by the Identity Problem), then the Asymmetry remains unsupported, and we are left with the possibility that there are reasons for bringing people into existence which are grounded on the prospective happiness of those people, and which are equally strong as the reasons against producing miserable people. But to concede this possibility would, of course, lead us straight back to the Total View.

VI. SUMNER AND THE AVERAGE VIEW

Although L. W. Sumner's essay is ostensibly "devoted to exploring and defending" the Total View, it in fact constitutes an attempt to refute the Average View. These two aims are in a sense complementary, for "since the average theory was devised precisely because it was thought to fare better than the total theory with problems of population" (p. 97), the defeat of the Average View in connection with these problems would provide an indirect and partial vindication of the Total View. Sumner's strategy is thus "to subvert the average theory by stages, first by removing one of its principal supports, next by showing that it violates one of the basic values of any utilitarian theory, and finally by outlining its singularly implausible implications" (p. 99). I shall review Sumner's arguments in the order in which he presents them, using his discussion as an introduction to an independent examination of some problems inherent in the Average View.

The prop of which, in his first argument, Sumner wishes to deprive the Average View consists in the claim, which he attributes to Rawls, "that the only version of utilitarianism obtainable by his [Rawls's] contract method is the average theory . . . " (p. 100). By demonstrating that, through the manipulation of certain features of the contractual situation, one can in fact supply a contractual derivation of the Total View, Sumner succeeds in restoring parity between the two views vis-à-vis contract theory (though both derivations rely on implausible assumptions).

Among the reasons which Rawls cites why the persons in the original position would prefer the Average to the Total View is that the latter "entails that so long as the average utility per person falls slowly enough when the number of individuals increases, the population should be encouraged to grow indefinitely no matter how low the average has fallen."[16] As long as the contractors are allowed to know that they will

16. John Rawls, *A Theory of Justice* (Oxford: Clarendon Press, 1972), pp. 162–63. The consequence cited by Rawls is (roughly) what Parfit has called the "repugnant conclusion" (see Secs. VII and VIII). Presumably, when Rawls speaks of the average falling "slowly enough," he means that it falls at an ever-slower rate so that it never drops below the zero point—that is, the point at which life ceases to be worth living. His statement of the consequence certainly requires this stipulation.

exist, it will be against their interests to select a theory which has this consequence. But, as Sumner recognizes, this consequence is not only objectionable from the self-interested point of view of the contractors; it is also morally counterintuitive. Sumner addresses this objection by conjecturing that in practice the Total View would not have this disturbing implication, but I find his reasoning overly sanguine and unconvincing (see Sec. VII).

Sumner's second argument charges that the Average View violates the requirement of "utilitarian impartiality" by discriminating against certain people simply on the basis of their "temporal location" (p. 103). He attempts to substantiate this allegation by providing an illustration. If we could produce a certain amount of utility either by increasing the utility of Smith, an existing person, or by bringing a new person, Jones, into existence, the Average View would require that we prefer the outcome in which Smith's utility is increased, even though the "gain" to each would be the same. The average would be lower if we were to produce Jones, because we would then have to divide through the same total of utility by a larger denominator. Sumner concludes that "the average theory favours Smith over Jones *simply because he already exists*. . . . The preference which the average theory exhibits is a pure (and arbitrary) preference for the living over the as yet unborn" (p. 103).

The reason that Sumner gives for condemning the "preference for the present over future persons" is that effects on people's utilities are "equally real regardless of where and when they occur" (pp. 102 and 104). We must ask, however, how "real" the effect on Jones's utility is if we "discriminate" against him. If we follow the dictate of the Average View and "give preference" to Smith, then Jones is not, as Sumner says he is, among those "affected by the choice"; he is not simply a person who is "as yet unborn"; he has no "temporal location." A fortiori, we have not discounted his utilities on the "morally irrelevant" ground of temporal location.

I think that Sumner's confusion here is the result of his falling victim to his own equivocal use of the phrase "future person." He often speaks of "future and contingent persons" (p. 103) or "future (possible)" persons (p. 109). In these instances he means by the phrase "future people" what Kavka does when he says, in his contribution to the volume, that "the trouble with future people . . . is not that they do not exist yet, it is that they might not exist at all" (p. 192). At other times, however, he talks, for example, of "future (as yet non-existent)" people, implying that these people will exist at some point in the future (p. 102).[17] It is this latter sense of the phrase that his argument requires, but the people against whom the Average View "discriminates" are of the type described by the former sense of the phrase.

17. Cf. Narveson's view that "future people are *real*. They differ from us by virtue of their location in time . . ." (p. 39). This seems to me the most felicitous use of the term. A future person is a person who will exist in the future. On the other hand, a person who might or might not exist in the future is a "possible person"; and a person who might exist or might have existed, but who in fact never will exist, is a "merely possible person."

The third argument which Sumner advances against the Average View does not, as he acknowledges, originate with him. It has, in fact, enjoyed a wide currency,[18] and therefore merits careful attention. The objection, stated crudely, is that the Average View would require us to be perpetually killing off below-average members of the population, for "every time we get rid of someone whose utility contribution is below average the average instantly rises (which means that still more persons are below it, whose elimination will raise it again, and so on . . .)" (p. 105). Sumner's exposition of this argument is instructive, since he explicitly states the assumption which is required to generate the objection:

> The average theory calculates its average over that pool of persons who will actually exist as a result of adopting a particular policy. . . . Thus if we choose to contract to a smaller pool, the average is calculated only over the members of that pool: the utility losses involved in eliminating existing persons or not producing new ones do not in themselves figure in the theory's calculations at all. [P. 105]

This claim can be challenged. As Derek Parfit notes, "It assumes that only the survivors matter. Why not say that everyone matters?"[19] Parfit's suggestion is that the relevant average is the average of all the people who ever live, not simply those who survive. The acceptance of this suggestion would entail that no one can be excluded from the class of persons whose average we are computing. Thus, even if an existing person is killed, he must nevertheless be counted in the denominator. Since, had he not been killed, his probable future utilities would have figured in the numerator, to kill him must, if his life would have been worth living, be counterproductive, since to do so involves lowering the numerator relative to the fixed denominator.[20] Therefore, on this understanding of the Average View, killing people can normally be expected to lower rather than raise the average.

This reply seems cogent. But, even though the Average View can thus be defended against this standardly invoked objection, there are, nevertheless, problems which remain. It might be urged that Parfit's revision (call it the "timeless version" of the Average View) manages to extricate the

18. Among those who have employed it are James E. Meade, Richard Henson, Robert Nozick, Philip E. Devine, Peter Singer, and another contributor to this volume, Robert Scott, Jr., (p. 78).

19. Parfit, "Overpopulation" (1973 draft—unpublished). Note that the understanding of the Average View which leaves the theory open to Sumner's objection *does* seem to violate the principle of utilitarian impartiality to which Sumner appeals in his second argument.

20. This claim assumes that the Average View employs a totaling principle in assessing single lives and applies the averaging principle only to populations. In other words, the data on individual utilities required by the Average View are being provided by the Total View. I assume that the application of the Average View to single lives (i.e., each person should aim for, say, the highest average per year lived) would be unacceptable; for it would entail that, if the remainder of a person's life were destined to be on average a little less good than his life has been hitherto, it would be better for his life to end now.

theory only at the cost of rendering it unworkable. In practice it would be impossible to calculate averages for all persons who ever live: we would be unable to fix the size of our denominator, and even if we could the utility information needed to fill out the numerator would remain inaccessible.

Perhaps this objection can be met by further revising the theory. It might be said that if we simply take the average for all those about whom we have reliable utility information, we will have not only a workable theory, but also a theory that is immune to Sumner's objection. For the people whom the theory (call it the "second revision") fails to consider would not, for that reason, be susceptible to "elimination" in the service of a higher average: since they never figure in the denominator, one could not get a smaller denominator by killing them. (Moreover, one does not, *ex hypothesi*, know whether or not they are below average.)

Perhaps we could simplify matters even further by stipulating that, in performing our calculations, we need only consider those among the living about whom we have reliable utility information. Since we cannot affect past people, this further restriction seems admissible. (Call this version of the theory the "third revision.")

Now in fact it makes a difference which of these revisions we choose.[21] To see that they are not extensionally equivalent we may have recourse to an example.[22] Suppose that we intend to add to the present population and that two alternatives are open to us. We could either produce a group with 100,000 members, all at utility level 5, or a group with 1,000 members, each of whom would have a utility level of 50. I shall call these the "A group" and the "B group," respectively. Suppose further that, among existing people, we have reliable utility information concerning 10,000 people. All of these have utility levels of 1, and their welfare will be unaffected by the choice between the A group and the B group. Finally, suppose that we also have reliable utility information concerning 1 million people who have lived in the recent past, and that these people, too, all had utility levels of 1. If we choose in accordance with the second revision, which requires us to consider all those concerning whom we have utility information (i.e., 1,010,000 people), the calculation will proceed as follows: if we add the A group (100,000 at 5) the average will be 1.36, whereas if we add the B group (1,000 at 50) the average will be 1.05; so we should opt for the A group. But, if we choose in accordance with the third revision, which instructs us to consider only those among existing people concerning whom we have utility information, the calculation will be different: if we add the A group the average will be 4.64, whereas if we add the B group the average will be 5.45; so we should add the B group. The two versions of the theory thus yield conflicting directives. The prob-

21. As Narveson is aware (p. 46).

22. My choice of figures here has been governed solely by the requirements of the example—otherwise the numbers are entirely arbitrary. The pretense to precision should not be taken seriously.

lem is that the average depends on the size of the "base" upon which one calculates.

The fact that averages are relative to the selection of the base also entails that the judgments of each of these revisions will be relative to the state of our present knowledge. Suppose we subscribe to the second revision. If we happen to be ignorant of the 1 million members of former generations, we will decide to produce the B group, but if we suddenly discover utility data on these 1 million people, we will have to reverse our decision in favor of the A group. An analogous point could be urged against the third revision with regard to our knowledge of the welfare of existing people.

Thus, the two revisions are not equivalent, and each contains an essential element of arbitrariness. It should be clear, moreover, that neither is equivalent to the timeless version, and therefore that neither can function as a substitute for it. The timeless version employs a different (though unspecifiable) base.

The timeless version alone escapes the charge of arbitrariness, in that there are no ad hoc restrictions on the group whose average it calculates, and the size of the base it employs thus does not fluctuate concomitantly with the extent of present knowledge. But, as we have seen, this version is unworkable: the required information is simply unobtainable. It might be maintained, however, that all consequentialist theories are afflicted with this problem to a greater or lesser extent, and that, in this case as in others, one must seek to approximate the right answers as nearly as possible. In other words, one might contend that the timeless version would provide the right answers were the requisite data accessible, and that the various ad hoc revisions can be regarded as workable yet imperfect substitutes yielding only approximately right answers.

This proposal might be interesting if the timeless version were plausible—but it is not. On this view, whether or not a person should have a child now depends in part on how well off—and how numerous—people were in the Stone Age, and it is difficult to see why this sort of consideration should be relevant. (Comparison between a child's welfare and the average level of his contemporaries might be thought relevant on grounds of equality, but inequality between people now and people in the remote past hardly seems to matter.) Thus, I conclude that even if the Average View can survive Sumner's assaults, its prospects do not seem very bright. But the extent to which the Total View might profit from its defeat remains to be seen.

VII. SIKORA AND THE TOTAL VIEW

R. I. Sikora's piece, which is probably the richest and most provocative contribution to the volume, seeks to establish a variety of points which are unified by the supposition that they all converge in support of the Total View. Among Sikora's numerous aims, there are two which I shall dis-

cuss. These are, first, to defend the Total View against certain objections, and, second, to provide an argument in its favor.

Sikora cites "two main objections" to the Total View. The first is what he calls the "possible persons objection": that the Total View assigns us obligations "relating to" or "concerning" possible persons. (pp. 147 and 115). He never makes explicit, however, what he takes the force of this objection to be. It seems to me that the only plausible interpretation of the objection is that it is absurd if a theory assigns us obligations which are owed to possible people. Sikora, in fact, states the objection in this form in an earlier paper.[23] But, as he states it here, it seems not to be an objection at all; for any theory could conceivably assign us an obligation "concerning" a possible person (e.g., could in some circumstance make it obligatory for us to have a child). Insofar as the objection takes the plausible form, moreover, it fails to hold against either the Total View or the Average View, since the obligations which these theories impose are not owed to anyone, whether actual or possible.[24] It seems, therefore, that the objection is misplaced, and thus we may ignore its role in Sikora's argument.

The second major objection to the Total View is "that it commits us to what Derek Parfit calls the repugnant conclusion: that if an extremely populous world with a happiness-average barely above neutrality would have a greater happiness-total than a much less populous world with a far higher happiness-average, we would be forced to choose the first" (p. 113). While he concedes that the Total View "cannot escape Parfit's objection," he brings forward a number of arguments in an attempt to diminish the force of the objection. The first of these arguments claims to show that most rival theories also entail the repugnant conclusion (henceforth the RC). The argument is based on the following hypothetical case. The world has been devastated by nuclear war and the quality of the survivors' lives has been reduced to a very low level. These remaining people, though sterile, can, at some cost to themselves, repopulate the world by rearing a great batch of "test-tube babies" who will live quite happy lives. If the survivors make the sacrifice, people will on average be much better off, though the survivors themselves will of course be worse off. Sikora observes that "any theory that takes it to be wrong to require those alive to make sacrifices to bring about the existence of possible persons must

23. R. I. Sikora, "Utilitarianism: The Classical Principle and the Average Principle," *Canadian Journal of Philosophy* 5 (1975): 409–19.

24. Sikora is, perhaps, misled by an ambiguity in the term "obligation." Many writers use the term in a wide sense in which it covers all moral reasons for acting. (Utilitarians in particular tend to use the term in this way, since their theory does not admit an especially rich range of categories of reasons for acting.) In this sense, obligations need not be owed. There is, however, a narrower sense of the term in which obligations form only one category of moral reasons. Obligations in this sense are owed to assignable individuals. Those writers who, like Schwartz and perhaps Narveson, seem to think both that obligations must be owed and that they exhaust all types of moral reasons are, presumably, sliding from the narrow to the wider use of the term.

choose . . . the world with a very low happiness level: that is, it must embrace the repugnant conclusion" (p. 115).

It is evident that something has gone wrong. We have been told that the Total View "cannot escape" the RC, but in this case the Total View conspicuously avoids the conclusion which Sikora regards as repugnant, since it would direct the survivors to raise the children. The reason that this anomaly has emerged is that the "RC" which appears in the case just cited fails to match the descriptions of the RC which Sikora gives elsewhere. The RC which is implied by the Total View, and with which Sikora is nominally concerned, consists in the claim that, for any world with a finite number of people, each of whom has a life of a very high quality, there is in principle a world which, because it is vastly more populous, is superior, even though each person in this world has a life which is only barely worth living.[25] The second world is deemed superior on the ground that its more numerous inhabitants together possess more of what makes life worth living. The assumption behind the conclusion is that losses in the quality of life can always be made up for by sufficiently large gains in quantity. But in the case which Sikora has constructed, the world which he claims it would be repugnant to choose does not have the greater total of utility; nor does it have the larger population. Thus, to fail to require the move from this world to the other is not to embrace the RC.

In spite of the fact that Sikora has not achieved his avowed aim, which is to show that theories which do not require us to produce extra happy people entail the RC, his argument will still have force if we find it disturbing that these theories have the implication to which he calls attention. How damaging to these theories is the exposure of this implication?

It should be noted that Sikora's example has certain features that may be influencing our response in subtle ways. For example, if the survivors do not make the requisite sacrifices, the human race will perish. Many people may take the view that, while we should not normally be required to make sacrifices in order to ensure that more happy people will exist, we should be required to do so in certain condtions—for example, when the race would otherwise die out. The example would lose much of its purchase if the alternatives were (*a*) requiring sacrifices to secure the fairly immediate production of the populous, high-average world; and (*b*) allowing for the continuation of the low-average world and the gradual repopulation of the world without sacrifice.

Another feature of the case which may affect our intuitions is that, since the survivors have lives which are just barely worth living, their making the requisite sacrifices would render their lives no longer worth living. Thus, it is perhaps extravagant to claim that it is "repugnant" if a

25. This description of the RC (which follows Parfit's) differs from the one initially stated by Sikora in that here each person in the larger world has a life just barely worth living, whereas in Sikora's description the people in the larger world have lives which are on average just barely worth living. Thus, Sikora's statement allows for the possibility of inequality while the one I have given does not.

theory does not demand that these people reduce their lives to that level. Indeed, Sikora himself elsewhere notes with approval "our ordinary tendency to think that it is more appropriate to ask for sacrifices from those who are well-off than from those who are badly off" (p. 122). (Notice also that the theories Sikora is criticizing do not prohibit the survivors from making the sacrifices; they only fail to require them to. The Total View leaves no corresponding loophole for the avoidance of the RC.)

Finally, there is a contrast between the RC and the implication which Sikora draws which deserves mention. The Total View regards the more populous world of the RC as intrinsically superior to the world in which people are better off. But the theories against which Sikora's argument is directed do not claim relative superiority for the world with the lower average; they favor it because the move to the larger would would be worse for the people who exist. Person-affecting theories, which are among the prime targets of Sikora's argument, necessarily focus on process considerations, since they are concerned with the effects of processes, events, acts, etc., on people. They provide no criteria for the global evaluation of possible worlds considered in isolation from their particular histories.

Given these doubts about Sikora's example, I would contend that, while his argument is not without some force, it is clearly not as powerful an objection as the RC is. Sikora has not, therefore, succeeded in showing that rival theories are open to an objection which is equally grave as that which Parfit and Rawls have urged against the Total View.

Sikora's second line of defense against the charge that the Total View entails the RC is based on the contention that "the low-average alternative of the repugnant conclusion would not be worse than the actual world" (p. 117). Presumably we are intended to infer that, since the actual world is not repugnant, the RC must be "something we can live with after all" (p. 116).

This argument assumes that the world of the RC is like the actual world in containing natural or social inequality. The world of the RC as Parfit understands it does not have this feature. But this suggests that, if natural inequality is undesirable, and if Sikora is right about the average quality of life in the actual world, then the world of Parfit's RC is better than the actual world.

Is life in the actual world on average just barely worth living? The suggestion that it is strikes me as most implausible. This would mean that nearly all those persons (perhaps the majority of the world's inhabitants) who are less well off than the average person have lives which are not worth having (or even worth not having). Sikora is thus led to the ungenerous conclusion that these people are mistaken in attributing value to their own lives. He is led to embrace this conclusion because he accepts a hedonistic account of what makes a life worth living (see Sec. IV). He concedes, however, that if we accept a nonhedonistic understanding of the notion of a life worth living, it is then conceivable that a life could be on balance unhappy and still be worth living. Thus, it could be true that,

even if Sikora were right about the hedonic character of lives in the actual world, these lives would be on average well worth living, that is, better than the lives in the world of the RC.

Sikora's argument is, however, misdirected. Even if the world of the RC were no worse than the actual world, this would not defuse the objection to the Total View. Sikora seems to assume that at least part of what is supposedly repugnant about the RC is the nature of the populous world itself. This is suggested by his use of the phrase "the 'repugnant' average" (p. 123). By comparing this world with the actual world he is attempting to show that there is nothing inherently repugnant about it. But what is objectionable about the RC is not the populous world itself but the comparative evaluation of worlds—that is, the claim that the more populous world is superior to the world in which people are all better off. Let us assume that Sikora is right about the character of the actual world. Then suppose we compare an enormous world in which people have lives like those in the actual world with a world which is one-tenth this size, but in which people are almost ten times as well off. The analogue of the RC in this case would be the claim that the second world is worse. The force of the RC lies primarily in the comparative evaluation required by the Total View.[26]

Sikora's third line of defense consists in the claim that the Total View would in practice never require us to add vast numbers of new people to the population, thereby diminishing the quality of existing people's lives.[27] In this he agrees with Sumner. I shall not discuss his arguments for this claim, but shall make only the following observation. If his reasoning were correct, and we could not in the world as it is increase the total of utility by increasing our numbers, this would not seem to be a cause for relief to a Utilitarian, but should seem as lamentable as the fact that we are limited in our capacity to increase the total of utility by increasing the quality of individual lives.

Having now examined Sikora's various indirect arguments for the Total View, I should like to comment briefly on the positive argument which he gives in support of this theory. The argument is as follows: For any large group of people we might add to the world, it is virtually certain that some members of the group will have wretched lives. This fact counts against the production of the group. Thus, unless there is positive value in adding happy people to the world, we must regard it as wrong, other things being equal, to produce the group. Sikora concludes that there must

26. There is, thus, more to the RC than the simple claim that the the Total View would require that we increase the population of the actual world at the expense of existing people. It is for this reason that the constructive proposal of Robert Scott, Jr.'s essay is unsuccessful. Scott claims (pp. 84–86) that by attaching to the Total View a principle that discounts the utilities of future people, we would be able to avoid the major objections to that theory. But the incorporation of a discount rate for the future will do nothing to prevent the Total View from yielding implausible judgements on the comparative value of possible worlds.

27. If his previous claim about the character of the actual world were correct, this auxilary claim would follow as a matter of logic.

be positive value in producing happy people—otherwise it would in practice "always be wrong *per se* to bring any large group of people into existence"; and this is an "absurd conclusion" (pp. 136 and 161).

To the extent that Sikora's reductio depends on what we would be committed to in practice, it is unsuccessful. For in practice other things are not equal, and there will usually be "something adequate to offset the wrongness of bringing the wretched people into existence" other than the potential happiness of the other members of the group—for example, the desire to have children on the part of the various potential parents.

The problem with Sikora's reasoning throughout this argument is that he tends to let the ceteris paribus clause slip out of sight (e.g., p. 141). If we keep this clause fully in view, Sikora's argument seems somewhat less compelling, for then the only cases in which the argument is applicable are those in which existing people fail, on balance, to derive any benefit from the addition of the new group—for instance, people might simply be indifferent. In order to test for the response which Sikora is seeking to elicit, we must be careful to screen out extraneous considerations such as the alleged right to procreate, which might override the badness of the outcome in which miserable people come to exist. When factors of this sort have all been excluded, not everyone's intuitions may be expected to favor Sikora's view. Still his argument does present a challenge.[28] For my part, I do find it implausible to suppose that it would be wrong to bring into existence a large group of people, consisting mainly of happy ones, when effects on existing people balance out. The problem is to reconcile this view with the repudiation of the RC.

VIII. GLOVER'S SUGGESTION AND THE MERE ADDITION ARGUMENT

Can we reconcile the view that there is value in the creation of people with lives worth living with the repudiation of the RC? One promising line of thought has been suggested by Jonathan Glover. Glover writes that "we may decide that we value people's lives having various qualities (which would put them high on the scale of 'worth-while life') and that the absence of these qualities cannot be compensated for by any numbers of extra worthwhile lives without them." He argues further that this view

28. There are variants of Sikora's argument which may be even more compelling. Here are two examples. First, any child one might have would have some negative utility in his life, and this counts against conceiving him. Thus, unless his prospective positive utility has positive value, and so is capable of canceling out the negative value of his negative utility, then it must, other things being equal, be wrong to conceive him. But if a child's positive utility can count as as counterbalance to his negative utility, then it should also count as a reason in favor of conceiving him. Second, there is some small probability that any child one might have would have a life that would be worth not living, and this counts against having children. Thus, unless this possibility is outweighed by the much greater probability that one's child would have a life worth living, then it must, other things being equal, be wrong to have children. But if the probability that one's child's life would be worth living can count against the possibility that it would not, then it should also count as a reason in favor of conceiving him.

draws support from an "analogy with common attitudes to what is valued within a single life."[29] There are, for example, some things which we value, though no amount of these things could make up for the loss of certain other things which we value more.

This argument appears to abandon the assumption of the Total View that what matters is the overall sum of utility (or whatever it is that makes life valuable). For if lives without the qualities Glover refers to have some positive value, and the value of lives with these qualities is not infinite, and if the two types of life can be ranked on the same scale, then there must be some number of lives of the former sort which together contain more of what is of value than a given number of lives of the latter sort.

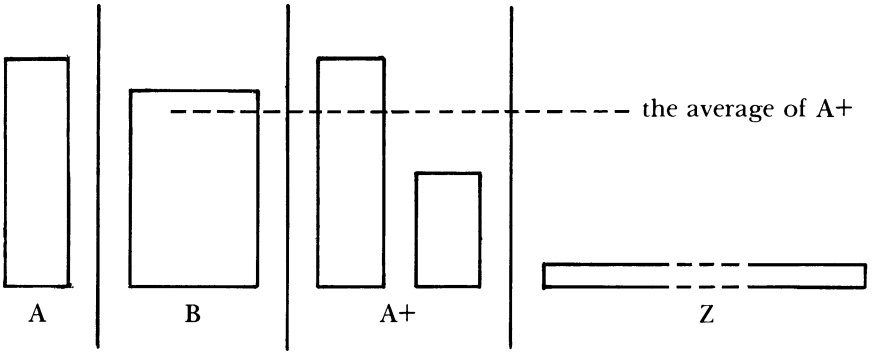
On the other hand, it might be argued that, if we pursue the analogy with values within a single life, Glover's theory might be rendered compatible with the Total View. Suppose we accept an analysis of utility in terms of the satisfaction of preferences. Suppose further that there are some things (call them "ϕ's") which I would rather have than not have, and which thus have some positive value, and that there are certain other things (call them "ψ's") which I value more than ϕ's, even to the extent that I would prefer to have a single ψ rather than any number of ϕ's. If what gives me the most utility is what I most prefer (given deliberative rationality), then we may be led to the paradoxical conclusion that the positive values of any number of ϕ's cannot be summed in such a way as to yield a total as large as the value of a single ψ. (Of course, this would not be paradoxical if the value of ϕ's diminished concomitantly with the increase in their number, or if the ϕ's were spread over time and I had a discount rate for the future; so we need to stipulate, perhaps implausibly, that neither of these conditions obtains.) One might then attempt to extrapolate this result to the interpersonal sphere by claiming that, by analogy, the value of certain lives is such that no number of these values can be summed to produce a total as large as the value of a single life of a different ("higher") sort.

There are objections to this latter argument, though space does not permit their full development here. I shall state them only briefly and without elaboration. First, the claim that what a person most wants is what gives him the most utility can be challenged, even from within the preference theory itself. What a person most wants, for example, may be a course of events in which far fewer of his desires are satisfied than might be in some alternative course of events. He may prefer the former course on, say, perfectionist or aesthetic grounds. To make it true by definition that the satisfaction of this single preference is what will maximize the individual's utility is to beg a number of important substantive questions. Second, it is not clear that an argument based on the preferences of a single individual can be extended to cover interpersonal cases. In the interpersonal case, there is nothing analogous to the preference for a

29. Jonathan Glover, *Causing Death and Saving Lives* (Harmondsworth: Penguin Books, 1977), p. 71.

single ψ over any number of ϕ's on which to base the claim that a single life might contain more of what is valuable than any number of lives of a lesser quality.[30]

It therefore seems better to regard Glover's suggested theory as an alternative to the Total View rather than (as I take Glover to imply) a version of it. Does this theory enable us to avoid the RC? Where arguments for the RC rely on the claim that what matters is the total amount of intrinsic value, or on the claim that there is value in the creation of extra lives that are worth living, I think it does. But Derek Parfit has formulated an argument, which he calls the Mere Addition Argument, which leads to the RC but which does not depend on either of these assumptions.[31] Consider the following diagram:



The rectangles designated by capital letters represent populations. While the horizontal dimension of each rectangle indicates the size of the population, the vertical dimension represents the level of welfare enjoyed (uniformly) by the members of the population. Thus, while each person in A is very well off, each person in Z has a life that is just barely worth living. The RC is simply the claim that Z, if large enough, is better than A. The Total View directly implies this claim.

But the same conclusion can be reached by a more circuitous route. This argument has two stages. The first consists in the claim that A+ is not worse than A. (It should be emphasized that the claim is not that A+ is better than A.) This seems unexceptionable, since the only difference between the two worlds is that A+ contains an extra group of people, all of whom have lives worth living and who cause no diminution in the level of welfare enjoyed by others. There are, however, two objections to this claim. The first is that A+ has a lower average. But, as Parfit points out, it seems absurd to object to a lower average "if it involves no loss of any kind, but mere addition." Indeed, the fact that the Average View judges

30. It would be interesting to explore the implications of a view of this sort for issues involving the saving of lives.

31. Parfit, "Overpopulation" (1973 draft), p. 4. See also his paper in Bayles, nn. 11 and 34.

A+ to be worse than A may be construed as a reductio against that theory. (This argument against the Average View is analogous to Sumner's third argument against that view.) The second reason why A+ might be considered worse than A is that it involves inequality. This feature may seem less objectionable, however, if it is stipulated that the inequality is not the product of an inegalitarian social system (e.g., the two groups may exist on opposite sides of the world and have no knowledge of each other). Moreover, when we are simply comparing A+ with A, the alternative to inequality is just for the worse-off group not to exist, so that, if anything, they benefit from the inequality. It therefore seems safe to conclude that A+ is not worse than A.

The second step in the argument consists in the claim that B is better than A+. This claim is supported by a number of considerations. For example, B has both a larger total and a higher average. The appeal to the average in this case, moreover, seems appropriate, since the same number of people are, on average, better off in B, which is also superior both in terms of the maximin rule and on grounds of equality. Again, the appeal to equality here seems more appropriate than in the preceding case, since in this case the avoidance of inequality does not involve the nonexistence of the worse-off group. (Indeed, it is possible, in comparing A+ with B, to assume that the people in B are the descendants of—or even the same people as—the people in A+, and that the level of the worse-off half in A+ has risen by more than the level of the better-off half has declined.)

Notice, however, that if B is better than A+, and A+ is not worse than A, then it follows that B is better than A. And the argument can be reapplied (B+ is not worse than B; C is better than B+ . . .) until we reach Z.[32]

To which step in this argument would Glover's theory object, and at what stage of the progression toward Z? The most plausible suggestion seems to be that the theory would object to the second step in the argument at the point at which the shift from the "plus world" to the world with the lower maximum quality of life would involve the loss of the valuable qualities to which Glover alludes. But, in this case, the sacrifice

32. Ronald Dworkin has recently pointed out that the claim that A+ is not worse than A need not entail that it is at least as good as A. There may be only partial comparability. This can be seen as follows: Suppose that there is another world, A++, which is like A+ except that the worse-off group in A++ is slightly better off than the worse-off group in A+. If we think that A++ is better than A+, and that A+ is not worse than A, then it seems to follow that A++ is better than A. But, in fact, we are not compelled to accept this conclusion. It may be that there is partial indeterminancy such that, while we may believe that each is neither better nor worse than A, it does not follow that each is exactly equally as good as A, nor that each is equally good as the other. If we accept Dworkin's claim that A and A+ may be only partially comparable, this will weaken Parfit's argument. The fact that B is better than something that is not worse than A will then not entail that B is better than A—only that it is not worse than A. The conclusion of the weakened argument will then be that Z is not worse than A. This conclusion is sufficiently counterintuitive for the argument to retain its force.

in the quality of life would not be made simply for the sake of increasing the number of lives. Thus, it is not clear that the objection actually follows from Glover's theory. In order to arrest the progression toward Z, Glover's theory would have to insist on the preservation of a certain quality of life for some people at the expense not only of total utility, but of average utility, maximin, and equality as well. If the level of quality it insists on preserving is high, then the theory will seem implausible in the face of this formidable array of intuitions. If, on the other hand, the level of quality defended is relatively low, the theory must allow that for any world whose inhabitants are all extremely well off, there is, in principle, some much larger world which would be superior, even though each of its inhabitants would have a life of this low quality. And this is, of course, a version, albeit a weakened one, of the RC.

It should be mentioned, in concluding this section, that Parfit's Mere Addition Argument is not an ad hominem against Glover. Since it shows that three perhaps equally compelling intuitions are incompatible, it poses quite a general problem.

IX. RIGHTS

Most of the preceding discussion has taken place within a broadly utilitarian framework. It is sometimes suggested that the fact that utilitarian theories encounter so many difficulties in the area of population theory simply underlines their inadequacy, while even the most recalcitrant population problems can be solved by an appeal to the theory of rights. It is held, for example, that the recognition of a right to a certain minimum level of welfare can provide a means of escape not only from the RC, but also from the disturbing implications of the Identity Problem—for example, that social policies which reduce the quality of life in the further future cannot be criticized on the ground that they are worse for future people. For if the prevailing level of welfare in the world of the RC, or in the world of the future if we implement certain policies, is lower than that to which people have a right, then these states of the world and the policies which bring them about can be condemned as involving the violation of rights. (We could also object to the move from A to A+ if the worse-off group in A+ would not be above the required minimum.) Finally, it seems that the theory of rights could account for what I earlier referred to as "the Asymmetry."

An initial response to these claims is to point out that the appeal to rights, even if otherwise successful, would not entirely solve either of the two problems mentioned. For, as in the case of Glover's theory, the problem of the RC could reemerge, albeit less disconcertingly, at a higher level. And the Identity Problem also reappears. If there is a choice between two policies, one of which requires minor sacrifices now in order to maintain a high quality of life in the future, while the other provides minor benefits now at the cost of lowering the future quality of life to just above the stated minimum, no one will be worse off, and no one's rights will be violated, if we choose the latter.

In spite of the fact that the argument from rights has these limitations, it would nevertheless, if successful, go far enough in the direction of a solution to warrant a certain amount of optimism. I shall argue, however, that it is unsuccessful. The objections to the argument are of two sorts, formal and substantive. I shall begin with the formal objection.

In the problematic cases with which the theory of rights is intended to deal, the persons whose rights would allegedly be violated can, in fact, exist only in a state in which their rights are violated. So the alternative to violating these people's rights is, in effect, to deprive the rights of their bearers. Thus, it would seem that we must say either that in preventing these people from existing we would be honoring rights without bearers, or that the rights which these people have can exist only when they are violated—in other words, that these people possess rights which can be violated but not honored.[33]

While I think that this formal objection can be met, the issues here are too complex to be adequately discussed in so short a space. I shall, therefore, simply draw attention to the fact that this problem has parallels both in this and other areas. If, for example, we can harm people by bringing them into existence, then we may have reason not to do what would harm people, even if we thereby ensure that there is no one whom we have spared from harm. Analogously, it might be argued that there is reason to avoid violating people's rights, even if one does not thereby honor anyone's right.

The problem we are considering is not peculiar to population theory. There are numerous cases in which our behavior is controlled by a regard for rights which people might, but do not have—for example, when one avoids entering into certain relations with a person because one would be unable to honor the rights which the person would then acquire against one. In this case it would be misleading to object that, in refraining from entering into these relations, one cannot be satisfying the constraints laid down by a right, since no right exists until these relations are established. The phrase "there is a right" or "a right exists" is ambiguous. On the one hand there is a restricted sense of the phrase according to which the claim that there is a right can be true only if there is an identifiable individual who possesses the right and to whom our obligations are owed. On the other hand, however, there is a broader sense of the phrase such that the claim that there is a right can be true even if there is no one who actually possesses that right. Given this distinction, there is no incoherence in the claim that there are cases in which one can respect a right by ensuring that no one (whether an actual or a possible person) comes to possess that right.[34]

33. This problem need not always arise in the case of the RC. If we were to make a rapid transition to the world of the RC we would violate the rights of existing persons—rights which could be honored.

34. In a case of this sort, the duty which is correlative with the right is not a duty which is owed to anyone. But population theory is an area in which we may expect the collapse of certain standard assumptions.

Even if, however, the formal objection can be met, there remain substantive objections which are not so easily evaded. The weakest of these points is the claim, standardly urged against the similar theories of the "social minimum" or "utility floor," that the specification of the acceptable minimum level of welfare must be to some extent both arbitrary and socially relative. I shall not dwell on this point, but shall simply note that, if the appeal to rights is to have any force against the RC and the Identity Problem, the minimum level of welfare to which people allegedly have a right will have to be well above the level at which life is just barely worth living.

The second point is perhaps more forceful. In the cases with which we are concerned, there exist people who have lives which are worth living, but which are nevertheless below the level of welfare to which people have a right. The theory of rights judges that it is objectionable in itself that these people should exist in this condition. But is this plausible? As we have seen, what is objectionable about the RC is the choice of the more populous world when there is an alternative in which people are better off—not the mere fact that in the world of the RC there exist people with low levels of welfare. Again, consider the choice between social policies mentioned earlier. What makes it wrong to choose the policy which lowers the quality of life in the future is connected with the fact that there is an alternative policy which would not lower the future quality of life. But if what is objectionable in these cases is not the mere fact that people with low levels of welfare come to exist, but is instead the fact that these people are brought into existence rather than some other group of people, then the appeal to rights is giving the right answers in these cases for the wrong reasons.

Is it, in fact, objectionable to conceive people whose lives, while worth living, would nevertheless be below some minimum level, even if it is not possible to produce happier people instead, and even if the existence of these people would not affect existing people for the worse? The affirmative answer given by rights theory may seem counterintuitive. If we believe that people can benefit by being conceived, then it will be in a person's interests to have received a marginally subminimal life (assuming that it is the only life he could have), even though this (allegedly) violates his rights. It is normally supposed that the primary function of rights is to protect people's interests; but in the present case the right in question has the opposite effect. (If we reject the view that conception can be a benefit, then the most that can be claimed is that to violate a person's right by bringing him into existence at a level below the minimum but above zero is not to act against his interests.) If those writers are correct who suggest that there is a conceptual connection between the ascription of rights and the protection of interests, then the objection I am pressing will be not only substantive but conceptual as well.

There are, of course, cases in which it would be wrong to violate a person's rights in order to benefit him. For example, it might be wrong to

violate a person's right to be told the truth even when it would be painful for him to know the truth. But cases of this sort are relevantly different in that the right in question will not be a right which is grounded on considerations of welfare. It will, instead, be concerned with the agent's dignity, autonomy, or something of this sort. While it may be plausible in these cases to suppose that considerations of dignity or autonomy override considerations of welfare, it is surely less plausible to suggest that we should refrain from benefiting a person simply in order to avoid violating one of his welfare rights.

Finally, it is arguable that conceiving a person with a life below the minimum but above zero would not constitute a violation of his rights at all. For, if the person is later glad that he was conceived, this might be regarded as tantamount to his retroactively waiving his right. If it is reasonable to assume that this tacit waiving will occur, then it is to that extent reasonable to assume that no violation of rights occurs when the person is conceived.[35] This claim is supported by an analogy with our treatment of young children. Suppose that, in order to save a child's life, we must perform an operation on him which will cause him to be seriously handicapped. It is the expectation that he will later be glad that the operation was performed that permits us to believe that we do not violate his rights by performing the operation without his consent.

For these various reasons I think that the appeal to rights, while perhaps formally admissible, is nevertheless implausible. There is a tendency in recent moral philosophy to try to couch discussions of all moral issues in the terminology of rights theory, but the couch can be a Procrustean bed.

X. CONCLUSION

The foregoing review of some of the major problems of population theory should have revealed how profoundly difficult these problems are. I suspect that one reason for this is that the intuitions that we have in this area seem, as it were, to have been borrowed from other areas rather than to have been conditioned directly by our exposure to the problems themselves. Until rather recently these problems were not sufficiently urgent at the practical level to intrude themselves upon our notice: people in the past had little capacity designedly to affect the welfare of future generations; they rarely contemplated the possible extinction of the human species by human action—and so on. Thus, it should not be surprising if it turns out that these problems bring to light certain hitherto submerged conflicts among our commonsense intuitions and, therefore, that no consistent theory will be without unappealing features.

35. One could also argue for this conclusion on the basis of the claim "that, in general, violation of an individual's right to something involves frustrating the corresponding desire." See Michael Tooley, "A Defense of Abortion and Infanticide," in *The Problem of Abortion*, ed. Joel Feinberg (Belmont, Calif.: Wadsworth Publishing Co., 1973), p. 60.