CLARK WOLF

O REPUGNANCE, WHERE IS THY STING?

On the Neutral Level of Existence and Parfit's Repugnant Conclusion

(from Tannjo and Ryberg, *The Repugnant Conclusion, Kluwer* 2004 pp. 61-80)

In philosophical discussions of ethics one often hears the complaint that arguments rely too heavily on intuitions. Of course intuitions also play a similar in other areas of philosophy: A system of modal logic may be rejected on the ground that it implies a counterintuitive view of counterfactuals; theories of justification may be rejected on ground that they have counterintuitive implications for knowledge, and theories of meaning may be rejected when they are judged to have intuitively unacceptable implications for a theory of reference. But intuitions are not a reliable guide to truth, and there is a point to the objection that they should not be used in the wrong way. As Derek Parfit has shown, many of our intuitions about the value of well being turn out to be confusing and contradictory when we apply them to populations of varying size and constituency. There is no normative theory of population choice that does not have seriously counterintuitive implications.

While intuitive judgments may not be a reliable guide to truth, our underlying intuitions will usually determine which view (if any) it is rational for us to accept. The fact that some moral theory would imply that American slavery was morally correct, for example, would be quite an excellent reason for most of us to reject that theory: the conclusion is so counterintuitive that we (most of us, as I hope) are fully justified in rejecting any philosophical theory that led us to it. Philosophical discussions about population choice often proceed in terms of intuition in just the way described above. Parfit's "repugnant conclusion" argument is just this sort: the conclusion is taken to be sufficiently repugnant to our evaluative and moral intuitions that we are justified in rejecting the plausible seeming premises that led us to it. But when philosophical arguments involve appeals to intuition, a great deal depends on the starting point and on which intuitions are held constant and which are allowed to float. I propose to examine the problem of population choice by starting with a plausible principle that most writers on this subject have been driven to reject. I must confess that I find the principle plausible, and find its rejection radically counterintuitive. This leads me to reject theories of population choice that would imply its denial, and to investigate the implications this principle would have for a broader theory of population choice. But

this may in part be a reflection on myself and on what, given my pre-theoretic intuitions, it is rational for me to accept. It may have nothing to do with the truth or falsity of the theory I investigate here. Still, I have found that my own intuitions on this subject are not unique, but are somewhat widely shared. This has encouraged me and led me to develop them further. It must also be noted that the view I develop here has counterintuitive implications of its own. I will do my best to identify and explain some of these implications, recognizing that some readers will regard them as objections. I hope that the exercise will be valuable for those who accept the argument as a *modus tollens* rejection of the premise as well as for those who, like myself, are driven to accept the conclusion. In this domain all theories come with hidden costs to our intuitive values, so this disadvantage is not unique to the view presented here.

1. COMPENSATION AND THE MISERY PRINCIPLE

Policy Choice: Suppose we are faced a population of six billion people who, through not fault of their own, are miserable and destitute. We have two policy options available to us. For the sake of the example, assume that these are the only alternatives, that they are described to include all the people who would ever exist, and that there are no other morally relevant effects of our choice.

> *Policy A* would improve the lives of all the miserable people so that their lives would be quite good. 'Policy A' would leave all them adequately well off with ample meaningful opportunities to pursue worthwhile projects.
>
> *Policy B* offers a very different opportunity. Instead of improving the lives of those who undeservedly suffer, we could instead choose a policy that would increase the number of well-off people in the world. Policy B would leave the lives of the miserable un-changed, but would spur a swift population explosion of many new people. These people's lives would be good, satisfying, and blissfully happy.

Some people believe that if the happy people who would come to exist under Policy B were numerous enough or happy enough, then their well-being would provide a counterweight for the misery of the unhappy six billion. For example, if we assume (as philosophers usually do and as economists usually do not) that utilities can be additively aggregated, then we might describe the choice above so that both total and average utility would be greater in Policy B. Even if we adopt a discount function that assigns less weight to the well being of "possible people" (as recommended by Partha Dasgupta (1989)), or one that assigns less weight to future well being (Arrow (1999)), or one that assigns a diminishing value to additional happy people as total population size increases

(Hurka (1983), Ng (1989), Broome (1992)), we can describe the choice above so that aggregate utility is larger if we choose Policy B instead of Policy A.

Many people share a powerful intuitive conviction that the choice of Policy B would be a moral mistake. If we wish to address the problem of human misery and destitution, we should provide those who suffer with opportunities, improve their lives, relieve them of their poverty, and bind their wounds. Bringing more happy people into the world is *no way at all* to deal with existing human misery. Derek Parfit articulates a principle that might lie behind this judgment. He calls it the *Claim about Compensation:*

> *Claim about Compensation:* One person's burdens cannot be *compensated* by benefits provided for someone else. (Parfit 1982, p. 337)

Parfit himself claims that "we cannot deny" the Claim about Compensation. Since many people seem to deny this principle, I am not convinced that one can't do it. But perhaps no one would deny the principle in the context where Parfit introduces it, where he is literally speaking of compensation. It is different in other contexts: in the context of theories of social choice over shifting populations, many economists and philosophers have adopted theories that imply the denial of the close relative of Parfit's Claim about Compensation that would apply to population and social choice. The Claim about Compensation's Social Choice cousin would urge that if there are people who are suffering poverty, deprivation, and misery, we cannot appropriately address this problem by adding more happy people to the world so that the happiness of the new people acts as a counterbalance to the misery of those who are suffering. I will call this the Misery Principle**.**

> *Misery Principle:* If people are badly off, suffering, or otherwise remediably miserable, it is not appropriate to address their ill-being by bringing more happy people into the world to counterbalance their disadvantage. We should instead improve the situation of those who are badly off.

Many people find the Misery Principle immediately plausible. But writers on population and social choice have mostly been led to theories that imply the rejection of this principle.

It is easy to see the relation between the Misery Principle and Parfit's Claim about Compensation. The existence of more happy people who would not have existed otherwise does not compensate those who are badly off, and from a broader social perspective we should not accept a social choice theory that urges us to be indifferent between policies that improve the situation of those who are disadvantaged, and alternative policies that would attempt to counterbalance their misery by encouraging fertility among the well-off. It's bad when people are miserable and deprived, and the

appropriate way to address such disadvantage is by improving the situation of those who suffer. In Charles Dickens' *Hard Times,* Sissy Jupe urges, "it must be just as hard upon those who were starved, whether the others were a million, or a million million."(Dickens, (1996, p. 94), Wolf (1996)) Just so.

Many of the considerations that make Parfit's *Claim about Compensation* immediately plausible can be cited in defense of the *Misery Principle* as well: Suffering and other relevantly similar burdens are a bad thing for those who endure them. Social outcomes that include uncompensated burdens are, other things being equal, worse than outcomes that do not include uncompensated burdens. Unless the existence of the additional well off people in Policy B somehow compensates for or counterbalances the unaddressed burdens of the miserable, Policy B is worse than Policy A. But the compensation principle tells us that the well being of people who would not have existed otherwise cannot compensate for unaddressed extant misery. Those who find it repugnant to purchase the happiness of many at the price of the misery of a few will find the choice of policy B repugnant for this reason.

We might list a set of important reasons for preferring Policy A over Policy B: Policy B involves persistent suffering while Policy A does not. The suffering of the miserable under Policy B is not compensated by the bliss of the happy. Finally, no one is worse off (or even badly off) if we choose Policy A, while some people are persistently badly off if we choose policy B. Surely it is better to have a world where everyone is well off, instead of an alternative world that contains unaddressed misery. And we can't change that ordering simply by adding more happy people to the world with miserable people in it!

But these reasons are not universally persuasive, since the Misery Principle rules out most of the widely defended theories of population choice. The theories that are ruled out by the misery principle include average, total, and number-dampened utilitarian views, as well as views that have been defended by John Broome (1992), Yew-Kwang Ng (1989), Partha Dasgupta (1989), Henry Sidgwick (1981/1907), Kenneth Arrow (1999), Nick Fotion (1997), Stuart Rachels (2001), and R.M. Hare (1993). Despite this daunting list of thoughtful and intelligent objectors, I would suggest that we should reject theories that are inconsistent with the *Misery Principle* unless, in the end, we find ourselves backed into a corner. I propose to consider what principles for population choice we reach when we hold constant the *Misery Principle* and its application in theories of social choice. I will show that there is a welfare theory, a version of utilitarian consequentialism in fact, that is consistent with the Compensation Principle in its broad interpretation as a restriction on social evaluation functions.

2. BETTERNESS, RIGHTNESS, AND THE ORDERING OF OUTCOMES

Philosophers have tools and language for discussing issues of choice and outcomes, but

the relevant terms are not always used in the same way by different theorists. It is therefore important to stipulate some terminology from the outset.

Consequentialism. Moral theories give us instructions about which actions we should choose, what outcomes we should regard as good, and what kinds of reasons for action should be regarded as *moral* reasons. A moral theory is *consequentialist* if it proposes that the consequences of our actions matter from the moral point of view. Consequentialists hold that outcomes should be evaluated in terms of their relative goodness, and that we act rightly, other things being equal, when we choose the action whose consequences are best.

Consequentialist theories may be *pure* or *impure*. Pure consequentialists hold that consequences are *all* that matter from the moral point of view. Many avowed utilitarians would seem to be pure consequentialists in this sense. Pure consequentialist theories hold that we should strive to discover which available actions or policies will have the consequences that are best, and should then choose one of the options in this best set. But pure consequentialist views have well known counterintuitive features: According to pure consequentialism, motives, principles, and special attachments to others are morally irrelevant except insofar as they result in morally desirable or undesirable consequences. According to pluralist consequentialist theories, consequences are *among* the things we should consider when choosing actions. But pluralists allow that we may take other considerations into account as well, and some stipulate principles or rights that permit or even require behavior that will produce consequences that are not always the best. While pure consequentialists may be monists about value and moral choice, pluralists may allow a variety of values that reasonably move us in different circumstances.

There are many possible ways in which nonconsequentialist considerations might influence our choices. For example, most people assume that it is permissible to give at least some priority to those we love when deciding what to do. At the margin, the "least impure" consequentialist theory would permit me to do this only when other things are equal: if my daughter and a stranger are drowning and I can save only one, then it is permissible for me to save my daughter instead of flipping a coin to give each swimmer an equal chance to be saved. But other pluralists hold that it is sometimes permissible or even obligatory to give nonconsequentialist considerations positive weight in deciding what to do: they might regard me as reprehensible if I were to flip a coin in the example above, or might permit me to give even stronger priority to those I love (for example, to feed and clothe my children even when people in other parts of the world are starving). Impure consequentialists might instruct us that there is reason to avoid lying or deception even when lying and deception would produce better consequences. And at the other extreme, a more strongly pluralist theory might even incorporate categorical requirements that trump consequentialist considerations in at least some circumstances. Such a theory could still be consequentialist, since other categorical values might not

always be at stake in all morally significant decisions.

Utilitarianism and Well-Being. A consequentialist theory is *utilitarian* if it ranks consequences in terms of the well or ill-being that they contain. Utilitarians are distinguished from other consequentialists by their insistence that it is well-being and ill-being that make outcomes good (or bad). But different utilitarians have different views about what constitutes well(ill)-being. I will use the paired terms 'well-being and ill-being,' happiness and misery,' 'bliss and suffering,' as interchangeable ordered pairs. This will allow me, I hope, to avoid treating here the complex issues involved in defining a metric of well-and-ill-being. My remarks here should therefore allow me to be agnostic about the scope of utilitarian concern (Are animals included?), about just what it is that makes lives go well or ill, and about the nature of well and ill being themselves. Suffice it to say that the account given should be consistent with contemporary theories that regard well-being to be a function of resources, welfare, capabilities, or preference satisfaction.

Many strongly consequentialist utilitarians seem to regard utilitarian reasons as *scientific* in some way, and regard nonconsequentialist prerogatives or imperatives as mysterious. It is also worth noticing that utilitarians are often moral realists of a special kind. Perhaps this is because happiness and well being, along with misery and ill-being, are very plausibly regarded as real properties of those who enjoy them. And happiness and well-being are clearly good for those who enjoy them, while misery and ill-being are bad for those who suffer them. So when we promote happiness or mitigate suffering, we are promoting something that is *real*. But to express this view is to understand what is wrong with it: even if happiness and misery (well-being and ill-being) are real, it would in no way follow that they should take precedence over other less tangible considerations in our deliberations concerning morally significant decisions. Strong consequentialists must either explain why consequences should play the such a role in our deliberative judgments, and why they should be regarded as uniquely moral so that other considerations (honesty, loyalty…) are non-moral. For example, strong consequentialists must explain why it is unjustifiable, irrational, or criticizable for me to give presumptive priority to those I love, to value truth-telling for its own sake. Pure consequentialists would need to explain why I should not regard myself to have a *prima facie* obligation not to murder other people even when the consequences of murder would be at least as good as the consequences of forbearing from murder. Both of these projects are difficult and perhaps impossible. Neither of them has ever been accomplished.

Commensurabilism. Most people and most moral theories regard well and ill-being as relevant from the moral point of view, and hold that it is presumptively obligatory (or at least morally good) to promote the former and mitigate the latter. But most consequentialist moral *theories* take well and ill-being into account in a specific way: it is usually assumed that well being is *commensurable* will ill-being, and that the two can be weighed against one another in a grand aggregate. I will call this assumption

commensurabilism. In order to understand the commensurabilist assumption, we need to consider positive and negative utilitarian imperatives separately. The following principles state general utilitarian versions of these imperatives:

> *Positive Utilitarianism (PU):* Actions are presumptively obligatory to the extent that they promote well-being.
>
> *Negative Utilitarianism (NU):* Actions are presumptively obligatory to the extent that they mitigate ill-being.

Phrased in this way, PU implies a presumptive obligation to maximize well-being, since actions that fail to maximize will leave us with additional presumptive obligations. Similarly, NU implies a presumptive obligation to minimize ill-being. In the absence of competing principles that assign non-utilitarian obligations, these two principles alone would constitute a form of strong consequentialist utilitarianism.

3. POSITIVE AND NEGATIVE CONSEQUENTIALIST IMPERATIVES

Suppose that, like most people, we regard human well-being and ill-being as relevant from the moral point of view. Then it is not a matter of moral indifference for us whether our choices and actions result in more or less misery or well-being. What are we to do when the positive and negative requirements conflict-- when the best way to promote well-being turns out not to be the same as the best way to mitigate ill-being? Before we will be in a position to answer this question, we need to consider the difference between these two principles, and some of the situations in which they conflict with one another.

Notice first that the positive utilitarian requirement is insatiable-- no matter how much well-being there is, PU implies that it is presumptively obligatory to produce more. On the other hand, NU is satiable-- if ill-being were eliminated, then NU would imply no further obligation. This may seem like a good feature of the negative view, since utilitarian theories are commonly taken to task for being too demanding. But reflection shows that this advantage is less decisive than one might think. After all, the negative principle also imposes weighty obligations on us, and in a world like our own, NU probably implies, like PU, that we are obliged to continue to make sacrifices until we are reduced to the level of diminishing marginal utility-- the point where any additional sacrifice on our part would produce more misery than it would alleviate.

There is another reason to question whether the fact that NU is satiable is really a mark in its favor. For one good way to minimize misery would be painlessly to extinguish all conscious life from the universe. A world without beings capable of misery would be a world in which misery was maximally mitigated, so NU might be

taken to recommend that we exterminate the miserable instead of helping them. Jonathan Glover, who gives brief but thoughtful attention to the negative view, writes "Martians may be negative utilitarians, but if they are we may hope that they will not find us."(Glover, 1995)

A second difference is equally striking: unlike PU, NU recommends that it is presumptively obligatory to make people happy (if they're not), but not obligatory to make happy people. In common philosophical parlance, NU is a "person affecting" moral principle, while PU is not. Because it is a person affecting principle in this sense, NU implies that *Policy A* is better than *Policy B.*

PU implies the opposite. In fact, PU recommends increasing the number of people in the world whenever this would promote aggregate well-being. If adding happy people to the world will increase this aggregate, as many utilitarians seem to think it could, then PU instructs us, other things being equal, to maximize the number of well-off people. The 'other things' that must be equal, of course, include the ill-being of those who are badly off. On a standard interpretation of PU, PU instructs us to increase population to the point of marginal utility, at which point the benefits afforded by the existence of an additional well-off person would be balanced out (compensated?) by the disadvantages that person's existence would impose on others.

There is a third important difference between NU and PU. If some people are miserable while others are well-off, NU assigns an obligation to aid those who are miserable, even if we could instead dramatically improve the lives of those we are well-off. But PU would recommend aiding the miserable only if the amount of well-being provided for those who are badly off is greater than the well-being we could instead provide for the well-off. If the well-off were more efficient at converting our aid into bliss, then PU would advise us to ignore those who are badly off and focus our attention on aiding the well-off. The two views focus our attention on different groups of recipients: PU would sometimes recommend that we ignore those who are badly off and serve those who are the most efficient producers of aggregate well-being. NU instructs us to ignore those who are well-off and focus our attention on those who are badly off.

It is surprising that most utilitarians seem to regard the negative principle as redundant or counterintuitive. Perhaps this is because most utilitarians simply assume that well-being and misery are commensurable currencies. Under the commensurabilist assumption, one way to promote happiness will be to improve the situation of those who are badly off, so combined with PU, commensurabilism implies a presumptive obligation to mitigate ill-being. So those who accept both PU and Commensurabilism are likely to regard NU as redundant.

The case is not symmetrical with NU: if we add commensurabilism to NU, we arrive at a theory that implies obligations that are both weaker and substantively different. NU is weaker for the obvious reason that it implies no obligation to maximize or even increase well being if no one is badly off. As I pointed out above, negative

obligations are satiable. It is different because it yields different prescriptions in some situations

But in the writings of utilitarian philosophers and economists, it is difficult to find much reflective consideration of the commensurabilist assumption. Usually it is simply assumed without much thought that if well-being is what matters, then well-being and ill-being must somehow balance one another out. If we take a broad collection of his writings into account, Mill is an ambiguous utilitarian and a weak consequentialist at best, but in *Utilitarianism* he offers one of the most famous and pristine conflations of the positive and negative imperatives: Mill (1979/1861) tells us that "actions are right in proportion as they tend to promote happiness, wrong as they tend to promote the reverse of happiness," and that "By happiness is intended pleasure and the absence of pain; by unhappiness, pain and the privation of pleasure." Mill famously fails to recognize that these are two commands, not one. He offers little help where the positive and negative imperatives conflict.

4. KARL POPPER AGAINST COMMENSURABILISM

Commensurabilists treat suffering as negative happiness, ill-being as negative well-being. They hold that either one can balance out the other depending on the situation. So if we have two alternative policies, one of which involves improving the lives of those who are badly off, while the other involves improving the lives of people who are already well off, commensurabilists tell us that we should be indifferent between the two, as long as the aggregate ill-and-well-being is the same regardless of which policy we choose.

Karl Popper regarded this as a serious mistake. He writes that one serious problem with the Utilitarian formula 'Maximize Pleasure" is that it "assumes, in principle, a continuous pleasure-pain scale which allows us to treat degrees of pain as negative degrees of pleasure. But Popper precedes Parfit's Claim About Compensation when he writes that "from the moral point of view, pain cannot be outweighed by pleasure, and especially not one man's pain by another man's pleasure."(Popper, 1962, vol 1, p. 234-5)

For the most part, Popper regarded himself as a skeptic about normative theory. Because of this he may seem an unlikely authority to cite in this context. But surely what he claims here is plausible and fits well with some of the pretheoretic intuitions that many people bring to discussions of normative population theory. If people are suffering and badly off, their situation presents itself to us as a kind of presumptive claim-- other things being equal, we have an obligation to help if we can. If we fail to aid others, though we could costlessly have done so, we have made a moral mistake and are blameworthy for it. There is no similar urgency when we have opportunities to provide

benefits to people who are well-off. Once again, Popper expressed this view with admirable clarity. And he also understood that this sense of moral urgency can be captured best if we regard the negative utilitarian imperative to be the most important one from the moral point of view.

> ...all moral urgency has its basis in the urgency of suffering or pain. I suggest, for this reason, to replace the utilitarian formula 'Aim at the greatest amount of happiness for the greatest number' or briefly, 'Maximize happiness', by the formula 'The least amount of avoidable suffering for all', or briefly, 'Minimize suffering'. Such a simple formula can, I believe, be made one of the fundamental principles (admittedly not the only one) of public policy. (Popper 1962 vol i, p. 235)

This passage commits Popper to "impure" or pluralist consequentialism. The imperative "Minimize suffering," as he insists, cannot be the only fundamental principle of public policy because we can prevent people from suffering by killing them. There may be circumstances in which some people would be willing to accept this extreme solution (mercy killing), but it is a conclusive intuitive objection to Pure Negative Utilitarianism that it would recommend moral indifference between helping those who are badly off, or exterminating them. Perhaps it is because the negative principle cannot by itself rule out this extreme and very final solution to suffering that negative utilitarianism has seemed so unattractive to most philosophers. But the objection only applies to *pure* negative utilitarian views. Pluralist theories can recognize a wide variety of standard moral reasons in favor of a general prohibition against murder. Popper's sensible insistence that the injunction to "minimize suffering" cannot be the *only* fundamental principle of public policy commits him to pluralism.

Popper specifically argues that there is "no symmetry between suffering and happiness, or between pain and pleasure," (Popper 1962, p. 284) and stresses that there is a difference in "urgency" between suffering and happiness:

> ...human suffering makes a direct moral appeal, namely, the appeal for help, while there is no similar call to increase the happiness of a man who is doing well anyway. (A further criticism of the Utilitarian formula 'Maximize pleasure' is that it assumes, in principle, a continuous pleasure-pain scale which allows us to treat degrees of pain as negative degrees of pleasure. But, from the moral point of view, pain cannot be outweighed by pleasure, and especially not one man's pain by another man's pleasure. Instead of the greatest happiness for the greatest number, one should demand, more modestly, the least amount of avoidable suffering for all; and further, that unavoidable suffering -- such as hunger in times of an unavoidable shortage of food-- should be distributed as equally as possible. (...) It adds to clarity in the field of ethics if we formulate our demands negatively, i.e. if we demand the elimination of suffering rather than the promotion of happiness. (Popper, 1962, pp. 284-5. Note 2 of Ch. 9)

Popper's view does require us to compare one person's ill-being with another, and adds a vaguely plausible distributional principle: unavoidable suffering "should be distributed as equally as possible." A full theory would need to say more about just how such comparisons are to be made, and what to do when the goal of equal distribution comes in conflict, at the margin, with the imperative to minimize suffering. These are

important questions, but they cannot be resolved here. I hope that any plausible answer to them will be consistent with the view I sketch in what follows.

5. POPPERIAN THEORY OF OBLIGATION AND BENEFICENCE

Suppose we wish to develop a theory to evaluate social outcomes that is consistent with Popper's advice. Such a theory would take the Misery Principle as a fundamental requirement. This severely restricts the range of possible theories, but the remaining theories will reject commensurabilism and will give the negative imperative to minimize ill-being presumptive priority over the positive imperative to maximize well-being. As a first approximation, we might consider theories that order the positive and negative imperatives as follows:

> *1. Negative Principle of Obligation [NPO]:* Actions that reduce (or minimize) misery are *prima facie* obligatory.
>
> *2. Positive Principle of Beneficence [PPB]:* Actions are good if they increase well-being. Actions are better or less-good depending on the extent to which they promote well-being.

In a previous paper I called this an "Impure Consequentialist Theory of Obligation," or ICTO. (Wolf, 1997) But the name and acronym are too unappealing to continue to use. Since this view is similar to that recommended by Popper, I will identify it as a version of "Popperian consequentialism." This theory, or rather the family of theories that are consistent with this ordering of the positive and negative utilitarian imperatives, regard the negative principle as assigning *obligations* while the positive principle applies to beneficence. It is good and praiseworthy, but not obligatory to do what will promote the well being of those who are already well off. But if we fail to honor our *prima facie* obligation to mitigate suffering and ill-being, then we may rightly be chastised, reprimanded, or even punished, depending on the circumstances.

I have taken some steps toward defending this view elsewhere (Wolf 1997), but for now I will note some of its attractive features. First, Popperian consequentialism comports well with many of our pre-theoretic judgments. Most people regard it as at least presumptively obligatory to address the suffering of others, but would not regard themselves as blameworthy if they were to pass up a chance to improve the situation of other well-off people around them. We expect praise and appreciation when we benefit our well-off friends, but we do not deserve a reprimand when we pass by a chance to delight our well-off neighbor and instead do what will only please ourselves. Consistent with this view, we typically regard it as less important, from the moral point of view, to

improve the situation of others who are already very well off. There is *urgency,* as Popper points out, in responding to the appeals of those who are badly off and who need our aid. There is no similar urgency about helping out the well-off, even if they are exceptionally efficient at converting our aid into bliss. Most of us would feel no compunction about providing benefits to others who are even moderately hungry instead of feeding Nozick's Utility Monsters.(Nozick, 1974, p. 41)

Second, Popperian consequentialism accounts well for commonsense judgments about genesis choices-- choices that involve bringing people into existence. It is commonly judged that we have at least a *prima facie* obligation not to bring into the world people who will be miserable. So if I know that any child I conceive will suffer from Tay Sachs disease, then it is presumptively obligatory for me to avoid conceiving children. While we might not actually blame others who conceive children under questionable circumstances in which the children they conceive risk serious disease or disability, it seems plausible that this is because we regard people as having a right to make their own decisions in such cases. The circumstances surrounding reproductive choices are complex and private, and we are rarely if ever in a position to pass harsh judgment on the reproductive decisions of others.

But while we may regard it as (at least) presumptively obligatory not to conceive a child who will be miserable, we do not typically regard it as obligatory to conceive children whenever the children we conceive would have an excellent chance to enjoy a good life. At most, we regard it as permissible or (at most) "good" to have children who face good prospects for a happy life. No one has an obligation-- even a prima facie obligation-- to have children. These pre-theoretic judgments are shared by many, and should be abandoned only if we have good reason to abandon them. While we should not put the wrong kind of emphasis on pre-theoretic intuitions about our obligations, especially in discussions of population ethics, it is at least a good feature or Popperian Consequentialism that it confirms widely shared judgments in the cases identified here.

6. POPPER AND PARFIT ON THE REPUGNANT CONCLUSION

The positive utilitarian imperative instructs us to promote well-being. But this principle is open to many different interpretations. Should we promote well-being by improving the situation of well-off people who already exist? Or should we do so by increasing the number of well-off people in the world? Is the impersonal value associated with adding more people to the population subject to diminishing returns as the total size of the population increases? Each of these views has been articulately defended. But each is subject to difficulties. One of the most pressing of these difficulties is that many popular views imply one or another version of Parfit's repugnant conclusion:

> *The Repugnant Conclusion*: For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living. (Parfit 1982, p. 388)

This conclusion seems repugnant all right, and it is worthwhile to develop a theory that can avoid it. Popperian consequentialism does not entirely succeed in allowing us to avoid the repugnant conclusion, but perhaps it is possible to undermine the repugnance by re-framing the conclusion itself.

Consider the choice problem involved by a Popperian Consequentialist (call him Karl) engaged in making Parfitian population choices. Faced with a collection of policy options, each associated with a range of possible outcomes, Karl begins by putting the outcomes in order according to their worseness or betterness. Because he accepts the Misery Principle, Karl gives priority to considerations of ill-being, and so he begins by ranking outcomes in terms of the ill-being they involve: Outcomes with more ill-being are worse than those with less. The negative principle does not by itself provide any basis for preference ordering (or 'goodness' ordering) among outcomes that involve equal ill-being. For that reason, we can call these NU-Indifference Classes.

Unless we reject the Misery Principle, we are committed to this first step for dividing outcomes into NU-Indifference Classes. It has some counterintuitive implications. Consider the ordering of possible worlds: This ordering principle implies that no world is better, in the initial cut, than the empty world. Some people also find it counterintuitive to think that no amount of well-being can compensate for an incremental increase in ill-being (Rachels 1999, p. 166), but intuitions on this issue go both ways. Since the argument if this paper holds the misery principle constant, it is unequivocally committed to this implication, counterintuitive or not.

Once we have identified the set of worlds that minimize misery, we may then give our attention to the second principle which instructs us to maximize well-being. Here we find that all of the problems Parfit discusses arise at a new level. If we simply wish to maximize well-being, we might do so by adding more well-off people, or by improving the situation of those who already exist. And as Parfit points out, for any world containing a finite number of very well-off people, there is an alternative world with more well-being in which everyone's life "barely worth living" or, as I prefer to put it, in which everyone's life is "barely good at all." Interestingly, NU does not by itself imply Parfit's repugnant conclusion. But neither NU nor Popperian Consequentialism will rule it out either. Thus by itself the Popperian view gives us no reason to prefer small well-off populations to large marginally well-off populations. We need to consider whether this is a problem. In what follows, I argue that the objection is less pressing if we carefully consider what it means for people to be "marginally well-off" or to have

lives that are "barely worth living."

7. ON THE NEUTRAL LEVEL OF EXISTENCE: WHICH LIVES ARE "BARELY WORTH LIVING"?

In his statement of the Repugnant Conclusion, Parfit refers to the notion that some peoples lives are "barely worth living." It is worthwhile to consider what this might mean, and which lives (if any) qualify. The phrase "barely worth living" may be prejudicial, since many people share the view that a life describable in those terms would be a bad life. The Repugnant Conclusion applies to lives that are good, but barely good. Since I suspect that this confusion about lives that are "barely worth living" influence people's judgments about the Repugnant Conclusion, I will refer instead to the level of well-being at which a life is just barely a good life. This gives us an only slightly changed version of the repugnant conclusion:

> *The Slightly-Less Repugnant Conclusion*: For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are just barely good.

Popperian Consequentialism is not committed to the Repugnant Conclusion, even in this slightly revised version. The Popperian view says nothing about how to evaluate the choice among outcomes that involve no ill-being. Trivially, suffering is minimized in any such case. But the view does not rule out the Repugnant Conclusion either: since it says nothing, it won't tell us that it's better to have fewer better-off people than it is to have more less-well-off people. To discover whether this is a problem, we need to consider more specifically what it means to say that someone's life is "barely good."

Which lives are just barely good? John Broome has very plausibly suggested that the neutral level may be vague (Broome 1997). We have useful conceptual tools for dealing with vague predicates, and we can apply them here. With most vague predicates, there are clear cases where the predicate applies, and fuzzy cases where the predicate seems to fade out. Just so in this case: the vagueness of the neutral level can be seen in the vagueness of the predicate "good" as applied to lives of various descriptions. I will assume without argument that some lives are unambiguously good, and that other lives are unambiguously bad. But suppose we describe different lives and try to order them according to their relative goodness. We might hope to find lives that are "barely good" by trying to identify those lives that are close to the vague neutral level, but which are clearly above it. As we approach the neutral level from above, we arrive at some lives that might either be judged to be good, or which might instead lie within the vague zone of neutrality (not to be confused with the zone of vague

neutrality). To interpret the repugnant conclusion, we need to consider the life situation of people we would regard as being only marginally above the neutral level. But the vagueness of the neutral level makes it difficult to discern which lives might be just marginally above it. Here are a series of suggestions for where we might set the neutral level and identify those whose situation is only marginally above that level:

> *Proposal 1:* Life is marginally above the neutral level at the point where one is indifferent between continuing to live and committing suicide.
>
> *Proposal 2:* Life is marginally above the neutral level at the point where one is indifferent between continuing to live and present death.
>
> *Proposal 3*: Life is marginally above the neutral level if it is dull, uneventful, and ordinary. In such a life, nothing in life is pleasurable, though nothing is painful or undesirable either.

Consider Proposal 1. It has sometimes been suggested that those who are at the neutral level are those who just enough reason to go on living that they forbear from suicide... for now. The Stoics sometimes implied that suicide was an easy "off switch" for life, available to us at any time if the burdens of life become heavy enough that we wish to heave them aside. And Julian Simon suggests that life must be worth living for everyone who hasn't yet committed suicide. (Simon 1981)

It must first be noted that Proposal 1 is seriously ambiguous. People may have a wide variety of reasons for considering suicide, and only some of these have to do with their own level of well-being. For example, one might prefer suicide if the only alternative were to bankrupt one's children with exorbitant medical expenses. So we need to limit Proposal 1 to people whose reasons are entirely self regarding.

Aside from committed Stoics, people who are indifferent to suicide or who are barely well-enough off that they forbear from suicide, are typically people who are miserable, suffering, and very badly off indeed. Suicide is a last extreme measure, and if people don't commit suicide it may not be because they (marginally) prefer life, but because they fear death. A thought experiment may be helpful in considering whether such a life qualifies as a marginally good life: Suppose you knew that any child you conceived would live a life in which she was always just barely well enough off that she forbore suicide. Most people would regard this as a strong reason not to conceive a child. If our children are as badly off as this, we have a weighty obligation to do what we can to improve their very bad lives. So this cannot be the neutral level that we should consider as we decide how to interpret Parfit's Repugnant Conclusion.

Note that Proposal 2 is different from Proposal 1: To be indifferent to death is not necessarily to be indifferent to suicide. Those who consider suicide must have

reasons to want to die, while those who are indifferent to death simply have no desire to live. Perhaps this indicates that Proposal 2 sets a slightly higher threshold than Proposal 1. But Proposal 2 is little better, and suffers from the same ambiguity, since people's indifference between life and death may have quite a lot to do with their fear of death. Since those who fear death will not be indifferent to it unless life is very bad (bad enough to overcome their fear), such indifference is not evidence that one's life is at the neutral level. So a marginal preference for life does not demonstrate that one is marginally above the neutral level.

In other respects, people whose lives offer so little pleasure or fulfillment that they are indifferent to death are in a bad situation. Their lives are bad lives, and it is quite plausible to think that we would have a presumptive obligation to help them if we could. If we would ourselves regard it as a bad thing to live such a life, then we cannot judge that people whose lives leave them indifferent to death are leading lives that are good-- not even marginally good.

Proposal 3 seems to reflect the way many people characterize the neutral level: A neutral life would be a life of muzak and potatoes, a life marginally better would be a life that "gives slightly more pleasure than pain, but the pleasures themselves are bland and boring." (Fotion 1997, p. 95) But this won't do: once again, a boring life is a bad life, just as Muzak is bad music. If we know people whose lives are like Muzak, we would have a presumptive moral reason to improve their lives and make them better off. No wonder people regard the repugnant conclusion as repugnant if they think of it as inhabited by people who are all living lives like these!

I strongly suspect that much of the repugnance associated with the repugnant conclusion derives from the way in which marginally good lives have been characterized. We think of a world full of billions of people who live lives *like that*, and it seems blah and drab and bad. If we knew that any children we had would live lives *like that*, we would have a presumptive reason not to have children at all. But this is itself suggestive.

8. O REPUGNANCE, WHERE IS THY STING?

If you knew that any child of yours was doomed to slavery, or destitution, or to a short and painful life followed by a long painful death, then it is plausible to think that you would have a strong moral reason not to conceive a child. Those who accept Popperian Consequentialism can explain that this reason is generated by the Negative Principle of Obligation (NPO). On the other end, perhaps there is some level of well being that is sufficiently high that if you knew that any child of yours would enjoy at least that level of well-being, then consideration of your child's well-being would not be a reason against conceiving a child. If you were a Popperian, you would not regard yourself as having an *obligation* to conceive a well-off child, but you might regard it as a good thing to do so. Consider the lowest level of well-being that qualifies. Perhaps we should regard this as the point marginally above the neutral level?

> *Proposal 4:* Life is marginally above the neutral level just in case, if you knew that any child you conceived would enjoy that level of well-being, you would regard it as permissible or even good to conceive a child.

Correlative to this, we could define the neutral level of well-being as the level such that if you knew that any child you conceived would enjoy that level of well-being, consideration of your child's welfare would not give you a reason not to conceive a child. This proposal would set the neutral level quite a bit higher than the others considered in the previous section. There is a plausible argument in favor of this proposal: What we want for our children is that they should live good lives. What worries us is that they will suffer bad lives. The level of well-being that would (just barely) satisfy our concerns about such matters should be the lowest level at which we regard our child's life as good.

But our relationship with our children is special, and extraneous considerations may enter because of it. Many parents want their children to be *very* well off, and will not be satisfied with a guarantee that their child will enjoy a life that is only marginally good. So maybe when we consider the welfare of our own children, we will be inclined to set the Neutral Level inappropriately high. But perhaps we can revise this proposal in a way that addresses this problem of partiality:

> *Proposal 4A*: Life is marginally above the neutral level just in case if you knew that any child *someone else* conceived would enjoy at least that level of well-being, you would not regard that person as having a reason (deriving from consideration of the child's welfare) not to conceive a child.

As we consider Proposal 4A, it may be possible to abstract from the partiality we feel on behalf of our own children. And perhaps this is a fair measure of the limit at which life is good.

I confess that when I consider the world of Parfit's repugnant conclusion to be inhabited by people who are all as well off as I would have them in proposals 4 or 4A, the repugnance has worn thin. In fact I find it a stretch to regard this conclusion as repugnant at all: most of the problems we associate with large populations are Malthusian problems, associated with limited resources and the misery of deprivation that might come from dividing a finite pie too many times. Given world enough and time, large populations would not be a problem. If everyone in the repugnant conclusion were sufficiently well off, the repugnance would be gone. And if we regard a life as good—even marginally good—then we should not regard it as a bad thing if there are people enjoying life at that level. And by stipulation, a life just above the marginal level would be a good life. How are we to test whether we really regard such a life as good? We can test it against the repugnant conclusion itself. If we find the

repugnant conclusion repugnant, perhaps it *follows* that the lives we are considering aren't really marginally good after all. This suggests yet one more proposal for finding the Neutral Level:

> *Proposal 5*: At some level of well-being, it is no longer repugnant to think of numerous people or even innumerable people living at that level, no longer odd to think that we might reasonably be indifferent between a world in which fewer people were even more blissful, and a more numerous world in which people were only just so blissful. Life is marginally above the Neutral Level when it is at least as good as this.

By raising the Neutral Level, we make the Repugnant Conclusion less repugnant. But this advantage comes at a cost which we must now count.

9. FROM REPUGNANCE TO ABSURDITY

The preceding section may be interpreted as a proposal that the threshold for the Neutral Level (what Parfit calls it the Valueless Level) should be set very high. But Parfit recognizes that raising the Neutral Level makes the Repugnant Conclusion less repugnant. But he adds that this method for resolving the Repugnant Conclusion solution leads to yet another counterintuitive result:

> *The Absurd Conclusion*: Suppose that, in some history of the future, there would always be an enormous number of people, and for each one person who suffers and has a life that is not worth living, there would be ten billion people whose lives *are* worth living, though their quality of life is not quite as high as the Valueless Level. This would be *worse* than if there were no future people. (Parfit, 1982, p. 415)

Is this conclusion absurd? If someone has a bad life, that is a bad thing from the welfare perspective. But if some possible people who would have had happy lives simply fail to exist, that's not a bad thing on the Popperian view. Some people regard this as counterintuitive, and indeed some regard it as a conclusive reason for rejecting views of the sort I have described here. But if one thinks that the value of well-being has to do with its importance to people—that is, it's significance in the lives of those who enjoy or lack it-- then this conclusion will not seem so absurd after all. But if one views well-being as something that has impersonal value, such that there is more of this value when we heap up more well-being, then the absurd conclusion will seem absurd indeed. Popperians must accept the implication that the empty world is at least as good, from the perspective of minimizing misery, as any possible world. More counterintuitively

perhaps, the Popperian view ranks the empty world as better than any world in which there is at least one person whose life is marginally bad—the actual world, for example. Some people may find this counterintuitive since it would seem to cast aspersion on the world in which we find ourselves.

In the end, whether the repugnant conclusion is repugnant and whether the absurd conclusion is absurd will depend on how one understands these conclusions, and on the intuitive judgments one brings to these problems. So here the argument becomes personal once again: Because I find it quite counterintuitive to regard well-being as impersonally valuable, I do not find the Absurd Conclusion excessively absurd. But note that we may have other non-welfarist values that lead us to want the human race to continue, and it may be these values that make Parfit's conclusion seem absurd to us. We may think that an empty world would be missing something morally significant, that it would be empty and cold, or that it would fail to provide opportunities for nobility and virtue since there would be no one to be noble or virtuous. A pluralist about values will have no difficulty in acknowledging such values, and there is no obvious reason to think that those of us who have them are thereby irrational. But a monist will be saddled with the responsibility to show that these values must all be subjugated under a monistic principle. I know of no good argument for the view that we should abandon or subjugate our non-welfarist values under a simple principle. Perhaps the very idea that a simple welfarist moral theory would or should supplant or trump other values we rationally hold should seem odd indeed. Popper rightly insists that the Negative Utilitarian principle cannot be the *only* principle of public policy, implying that this principle should be qualified by other non-welfarist considerations. Unless we are inclined to accept the truly repugnant notion that we can counterbalance the misery of some with the well-being of others, we should find this response satisfactory.

REFERENCES

Arrow, Kenneth. “Discounting, Morality, and Gaming”. In Paul R. Portney and John P. Weyant, eds. *Discounting and Intergenerational Equity*. Washington DC: Resources for the Future, 1999.

Broome, J. *Counting the Cost of Global Warming*. Cambridge: White Horse Press, 1992.

Broome, J. "The Value of a Person". *The Proceedings of the Aristotelian Society*, supp. vol. LXVIII (1992).

Broome, J. “The Welfare Economics of Population”. *Oxford Economic Papers.* 48 (1996): 177-193.

Broome, J. “Is Incommensurability Vagueness?”. In Ruth Chang, ed. *Incommensurability, Incomparability, and Practical Reason.* Cambridge MA: Harvard University Press, 1997.

Dasgupta, P. “Population Size and the Quality of Life”. *Proceedings of the Aristotelian Society*. 58 (Suppl) (1999): 23-40.

Dasgupta, P. "Savings and Fertility: Ethical Issues". *Philosophy and Public Affairs*, 23 (1994): 99-127.

Dasgupta, P. *Human Well-Being and the Natural Environment.* New York: Oxford University Press, 2001.

Dickens, Charles. *Hard Times.* Toronto: Broadview Press, 1996.

Fehige, C. "A Pareto Principle for Possible People". In Christoph Fehige and Ulla Wessels eds., *Preferences*. de Gruyter: Berlin/New York, 1996.

Fotion, Nick. “Repugnant Thoughts about the Repugnant Conclusion Argument”. In Nick Fotion and Jan Heller, Eds., *Contingent Future Persons.* Dordrecht: Kluwer Academic Publishers, 1997.

Glover, Jonathan. “The Research Programme of Development Ethics.” In Martha Nussbaum and Jonathan Glover, eds. *Women, Culture, and Development.* New York: Oxford University Press, 1995.

Hare, R. *Essays on Bioethics.* New York: Oxford University Press, 1993.

Hurka, T. "Value and Population Size". *Ethics* 93 (1983):496-507.

Mill, J.S. *Utilitarianism*. Indianapolis: Hackett Publishing Co., 1979/1861.

Narveson, J. *Morality and Utility*. Baltimore MD: Johns Hopkins Press, 1967a.

Narveson, J. "Utilitarianism and New Generations". *Mind* 76 (1967b): 62-72.

Narveson, J. "Moral Problems of Population". *The Monist* 57 (1973): 62-86.

Ng, Y-K. "What Should We Do About Future Generations? The Impossibility of Parfit's Theory X", *Economics and Philosoph*, 5 (1989a): 235-253.

Nozick, R. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Parfit, D. *Reasons and Persons*. Oxford: Clarendon Press, 1984.

Popper, K. *The Open Society and its Enemies*. Princeton N.J.: Princeton University Press, 1962.

Rachels, Stuart. “Review Essay: Contingent Future Persons”. *Bioethics* 13(2) (1999): 160-167.

Rachels, Stuart. "A Set of Solutions to Parfit's Problems". *Nous* 35(2) (2001): 214-238

Rawls, J. *A Theory of Justice*. Boston: Harvard University Press, 1971.

Scanlon, T. "Preference and Urgency". *The Journal of Philosophy* 72(19) (1975): 655-669.

Julian Simon. *The Ultimate Resource*, Princeton: Princeton University Press, 1981.

Sidgwick, H. *The Methods of Ethics.* Indianapolis: Hackett Publishing Co., 1981/1907.

Singer, P. "A Utilitarian Population Principle", in Michael Bayles, ed. *Ethics and Population*, Cambridge: Schenkman Publishing Co., 1976.

Smart, R.N. "Negative Utilitarianism". *Mind* 67 (1958): 542-3.

Vetter, H. "Utilitarianism and New Generations". *Mind* 80 (1971): 301-2.

Wolf, C. "Social Choice and Normative Population Theory: A Person-Affecting Solution to Parfit's Mere Addition Paradox" *Philosophical Studies* 81(2-3) (1996): 263-282.

Wolf, C. “Person-Affecting Utilitarianism and Population Policy; Or, Sissy Jupe’s Theory of Social Choice.” In Nick Fotion and Jan Heller, Eds. *Contingent Future Persons*, Dordrecht: Kluwer Academic Publishers, 1997.