

3. Deterrence, utility, and rational choice

The fundamental question of this book is whether practicing nuclear deterrence, in any form, is morally permissible. The present chapter is an attempt to deal with this question from the point of view of utilitarian moral theory.¹ For reasons given in the Introduction, I believe that the sort of minimum deterrence policy sketched there is, from the utilitarian perspective, far superior to the present deterrence policies of the superpowers.² With policies of the current type eliminated from contention as the choice recommended by utilitarian considerations, this chapter seeks to discover the best utilitarian policy by comparing minimum deterrence with the alternative of not practicing nuclear deterrence at all, that is, unilateral nuclear disarmament. It poses the issue as a problem of rational choice under conditions of uncertainty, reveals difficulties with the expected utility and maximum approaches toward solving it, and proposes an alternative principle of choice that may plausibly be applied to achieve a solution.

I begin with some simplifying assumptions. (Whether these assumptions distort or bias the analysis will be discussed in section V of this chapter.) Only the bilateral superpower balance of terror will be considered; complications due to the existence of other nuclear powers are ignored. As noted above, attention will be limited to a superpower's choice – hereafter called the *deter-or-disarm* choice – between the basic alternatives of (i) unilateral nuclear disarmament,

A remote ancestor of this chapter was presented to the Jacob Marschak Interdisciplinary Colloquium on Mathematics in the Behavioral Sciences, at UCLA, in the spring of 1976. I am grateful to the members of that group, and to Robert M. Adams, for their helpful comments. My work was partly supported by a Regents' Faculty Research Fellowship from the University of California.

and (ii) practicing some form of minimum deterrence (of the kind described in the Introduction) that involves the possession of nuclear weapons and the threat of their use against one's rival's homeland. Finally, it is assumed that a nation is choosing a policy for the foreseeable future, which is to be thought of as a significant but limited period, say thirty years.³

Our question is what now ought to be done by each participant in the balance of terror, given that participant's perceptions of the present political-military situation. The crucial elements of each superpower's perceptions are, presumably, these. First, it believes that its ideological-economic system promotes the well-being of those living under it to a much greater extent than does its opponent's system. Second, it believes its rival would probably impose its system on other nations if it could. Third, it regards complete bilateral nuclear disarmament as unattainable in the near future, largely because of its rival's unwillingness to disarm on reasonable terms. Our question is which of two nuclear policies – minimum deterrence or unilateral disarmament – would a *utilitarian* sharing a superpower's perceptions of the balance of terror wish that superpower to follow. Alternatively, we could ask whether a superpower would practice minimum nuclear deterrence (rather than unilaterally disarming itself of nuclear weapons) *if it were to decide the matter purely on utilitarian moral grounds*.

The dilemma a superpower would face if it looked at the balance of terror in this way is clear. It would recognize that it could greatly reduce (and perhaps eliminate) the danger of large-scale nuclear war by disarming unilaterally. But it would fear that if it did so, nothing would stand in the way of its rival dominating the world by the use of, or threat to use, nuclear weapons. Such a result would, given its view of its opponent's system, constitute an immense utilitarian disaster. On the other hand, it would realize that its continued participation in the balance of terror runs the risk of resulting in a different utilitarian disaster – large-scale nuclear war. How can this utilitarian moral dilemma be resolved?⁴

I. UNCERTAINTY

The fundamental principle of utilitarianism is that agents should act to produce as much long-run net utility for people as possible,

where utility is conceived of as some objective quality such as pleasure or desire satisfaction.⁵ This principle is generally interpreted to imply that, when the consequences of one's acts are not known with certainty, one should perform the act having the greatest expected utility. We cannot, however, determine the utilitarian status of nuclear deterrence by simply comparing the expected utilities of deterrence and unilateral nuclear disarmament. For application of the Expected Utility Principle (EUP) to this case would require us to have something we lack: reliable quantitative utility and probability estimates.⁶

Consider probabilities first. A distinction is often made between choices under risk, when the probabilities of the various outcomes following from the acts in question are known, and choices under uncertainty, when they are not known. The deter-or-disarm choice must be made essentially under conditions of uncertainty. This choice presents itself in a complex and unique historical situation. Assigning numerical probabilities to the possible outcomes of the available choices would require selecting an appropriate reference class of past situations to provide data on the relative frequencies of various outcomes. But the results will depend entirely upon the reference class selected. The probability of deterrence leading to war, for example, will turn out rather high if we choose past arms races among major powers as our reference class, but will seem very low if we choose earlier segments of the present balance of terror as our reference class. Lacking a natural reference class, such as we have when estimating the probability of rolling a six on a throw of a symmetrical die, and lacking a scientific theory of international relations that would allow us to deduce numerical probabilities of outcomes in some indirect manner, we cannot expect reasonable people to agree on the probability that minimum deterrence would lead to war, or that unilateral disarmament would lead to domination.

Serious problems also arise if we attempt to compare quantitatively the utilities of such outcomes as "large-scale nuclear war" and "rival's world domination."⁷ In general, we face two problems in making quantitative estimates of utility: determining how utility is to be defined and measured, and actually carrying out the measurement process. Problems of the first sort aside, the practical difficulties of measurement in all but the simplest cases are so great that utility estimates must be based on crude simplifying assumptions

(e.g., that negative utility is proportional to the number of casualties). In the present case, added to these standard difficulties is the fact that the nature of the specified outcomes is unknown, so that their effects on people (and hence their utility) cannot be accurately estimated. Lacking empirical data on the effects of large-scale nuclear war and world domination by the opposing superpower, we are in a poor position to estimate quantitatively the utilities of these outcomes. Indeed, as hinted at above, the term "world domination" is really shorthand for a wide variety of outcomes, such as tighter or lesser control of other nations by one's rival brought about by nuclear blackmail (implicit or explicit), nuclear attack, or victorious nuclear war against lesser nuclear powers. Obviously, the amounts of negative utility produced by these different outcomes, like the amounts of negative utility produced by different forms of large-scale nuclear war, will vary a great deal. In the sequel, therefore, we should remember that the terms "large-scale nuclear war" and "rival's world domination" cover a range of more specific outcomes, and should think of the utilities of these general outcomes as rough averages of the utilities of those more specific outcomes.

Without reliable quantitative utilities and probabilities to work from, we either cannot apply EUP, or must do so using unreliable estimates (i.e., guesses) that we have little confidence in.⁸ The numbers emerging from such expected utility calculations can hardly form a sound basis of moral decision. Recognizing the problem of applying the traditional EUP under conditions of uncertainty, decision theorists have attempted to extend the scope of application of that principle by reinterpreting probabilities and utilities as entirely subjective.⁹ They show that if an agent supplies sufficient data about his preferences between lotteries having specified probabilities of various outcomes, and these preferences satisfy certain plausible axioms, then quantitative subjective utility and probability measures for the agent can be constructed such that (i) the agent's expressed preferences maximize expected utility, (according to the constructed measures), and (ii) maximizing expected utility in accordance with the constructed measures in other decisions involving the given outcomes is the only way that the agent can make these decisions consistent with his expressed preferences and the axioms.

While this result is a substantial contribution to rational decision theory, it cannot be used to solve the deter-or-disarm problem from the point of view of utilitarian moral theory. This is partly because,

when making choices of such vast utilitarian importance as the choice between deterring and disarming, consistency with other actual or possible decisions pales in significance in comparison with the problem of arriving at the proper decision in the case at hand. More important, applying subjective utility analysis to our problem presupposes the agent's ability to express meaningful preferences between lotteries involving such outcomes as large-scale nuclear war and world domination by the opposing superpower. Does she prefer a 10 percent chance of war to a 70 percent chance of domination? A utilitarian agent confronting these questions can only guess, and will have little, if any, confidence in these guesses. But the subjectivist analysis depends heavily on such guesses, though the agent herself is entirely unconvinced of their value as a basis for choice.

In essence, the subjective utility theorist and the utilitarian moralist differ in their conceptions of the problem that the latter faces.¹⁰ The subjective utility theorist sees the utilitarian's problem in choosing between deterrence and disarmament as simply one of clarifying and making consistent the utilitarian's own subjective preferences. But the utilitarian, who wants to produce as much well-being and prevent as much suffering as possible, regards himself as attempting to deal with objective moral values in the face of extreme factual ignorance. If he knew the amounts of well-being and suffering that would be produced by war and by domination and the probabilities of these outcomes, or if he had what he regarded as reliable and objective scientific estimates of these quantities, he would probably be willing to use the traditional EUP in making his choice. Unlike the subjective utility theorist, he regards such knowledge and such estimates as possible in principle. But he does not in fact have them. Nevertheless, given the importance of the problem, he feels called upon to make a reasonable judgment based upon what he does know.

What then does he know that he could base a choice on? Aware of the enormous power of nuclear weapons, he presumably knows that a large-scale nuclear war would almost certainly be a worse utilitarian disaster than domination of the world by the opposing superpower (even if that domination were brought about by means of a small nuclear war). He does not know enough to determine or estimate accurately how much worse large-scale nuclear war would be, but he may rightly have confidence in the ordinal judgment that it would be worse to some unknown extent. A similar ordinal judg-

ment may be made about the relevant (conditional) probabilities. While unable to determine how much more likely, he may be able to assume with confidence that it is more likely that unilateral disarmament would lead to domination than that practicing a policy of minimum deterrence for the relevant time period would lead to large-scale nuclear war. How can it be possible to arrive at such a judgment when we lack reliable quantitative estimates of the probabilities in question? The answer is that different methods of analyzing the same situation may yield differing numerical estimates, but nevertheless agree unanimously (or nearly so) in their ordering judgments. Thus, while different theories of international relations, or different experts making thoughtful overall judgments, or the choice of different plausible reference classes, may yield widely ranging numerical estimates of the (conditional) probability of war or domination by the other superpower, these methods may, on the whole, agree that one (in this case domination) is more probable than the other. This appears to be the way things stand with respect to expert opinion, as seen from the point of view of the United States.¹¹

Let us call a choice between alternatives, any of which may result in disaster, a choice between *potential disasters*. Further, let us say a choice is made under *two-dimensional uncertainty* if the chooser has no reliable quantitative estimates of the relevant utilities and probabilities, but has confidence in his judgment of their ordinal rankings. What I am suggesting is that we view the deter-or-disarm choice as a choice between potential disasters under two-dimensional uncertainty. A choice of this sort is easily made if there is one alternative that minimizes both the probability of disaster occurrence and the degree of disaster, should one occur. However, in the deter-or-disarm situation, neither alternative possesses both features. Here one must choose, under two-dimensional uncertainty, between a smaller risk of a graver disaster and a greater risk of a smaller disaster. We have seen reasons for doubting that EUP can solve this choice problem for the utilitarian. In the next section, the desirability of applying another popular principle of rational choice is explored.

II. MAXIMIN

Among the most favored principles of rational choice under uncertainty is the Maximin Principle (MMP). It prescribes selecting the

available alternative with the best worst outcome, that is, the one yielding the maximum minimum payoff. MMP is, in one respect, well suited for application to the deter-or-disarm problem: its use requires only ordinal rankings of the utilities of the various possible outcomes. On the other hand, as will be noted in Chapter 7, MMP, strictly interpreted, does not allow us to choose between minimum deterrence and unilateral disarmament. For it is *possible*, though highly unlikely, that unilateral disarmament could produce the worst outcome of all, namely human extinction. (This could happen, for example, by the rival power having a nuclear war with other powers that caused an extreme and long-lasting nuclear winter.) Thus, comparing deterrence and disarmament by their worst possible outcomes does not distinguish between them.

We could, however, modify MMP, so that it considers only outcomes that have a *significant probability* of following from the acts or policies in question. It is not implausible to assume that minimum deterrence has a significant probability of resulting in extinction (through large-scale nuclear war and its environmental effects) while unilateral disarmament does not. Given this assumption, a utilitarian following the modified MMP would favor unilateral disarmament over minimum deterrence, since the worst significantly probable outcome of the former (say, rival world domination secured by limited nuclear attack) is not as bad as the worst significantly probable outcome of the latter (say, large-scale nuclear war leading to extinction). To see if modified MMP can plausibly be applied in this way to solve our problem, it will be useful to look first at Rawls's use of MMP in his theory of justice.¹²

Rawls seeks to discover and justify principles of social justice by asking what principles governing major social institutions would be selected, in an original position of choice, by rational self-interested parties expecting to live under those institutions. Each party in the original position is assumed to know that the society in question will be subject to "moderate scarcity" of resources, but is assumed to lack knowledge of her own individual traits (e.g., abilities and goals) that could enable her to bias the principles in her own favor. Rawls contends that such parties so situated would follow MMP and attempt to assure for themselves the highest possible security level, by opting for social arrangements ensuring the best worst outcome. The resulting conception of justice is embodied in Rawls's *difference principle*, which says that inequalities in the distribution of primary

goods (i.e., those goods such as income and liberty that it is rational for a person to want no matter what particular things she may want) are allowable only if they work to the greatest benefit of the least advantaged members of society. Rawls cites three main reasons for thinking the parties would follow a maximin strategy.¹³ First, since there is no positive evidentiary basis for a party assigning a numerical probability to her ending up in a particular social class or position, EUP cannot reasonably be applied. Second, each party would know that she would receive an acceptable supply of primary goods under the difference principle, but she might not receive an acceptable supply if principles aimed at maximizing expected utility were chosen. Third, the parties do not substantially prefer the larger amounts of primary goods that they might obtain under other principles to the acceptable supply they would be assured of under the difference principle.

However, even if we accept this reasoning as sound, it would not follow that modified MMP yields the correct solution to the deter-or-disarm choice problem. For there are crucial differences between the deter-or-disarm choice situation and the situation faced by the parties in Rawls's original position, so that none of the three reasons Rawls offers for following a maximin strategy in the latter situation applies in the former situation. First, the deter-or-disarm situation is not one of complete uncertainty. There is at least one piece of probability data that is known: the ordinal ranking of the two relevant conditional probabilities. Second, while Rawls assumes moderate scarcity of resources so that one can be assured of receiving a supply of primary goods sufficient to sustain a minimally decent life if MMP is followed, there is no choice available in the deter-or-disarm situation that ensures an acceptable outcome. Third, in the deter-or-disarm situation, utilitarian choosers greatly prefer the favorable outcome they might achieve by following the nonmaximin policy (i.e., preservation of the status quo) to the security level outcome that could be assured by their playing a maximum strategy (i.e., world domination by the rival). To see how these differences affect the plausibility of applying MMP, let us alter Rawls's choice situation so that it comes to resemble the deter-or-disarm choice situation in these respects.

Let us suppose that there is some minimum amount X of basic primary goods (e.g., food and time free from labor), such that life for almost anyone receiving less than this amount would contain

miseries outweighing its joys and would not be worth living.¹⁴ *Extreme scarcity* may be said to exist in a society if there is no way of organizing production and distribution so as to yield a supply of basic goods equal to or larger than X to the members of the worst-off class. Imagine that the parties in the original position know that their society will be operating under conditions of extreme scarcity. If primary goods are distributed according to the difference principle, a substantial minority will receive an allotment of basic primary goods that is somewhat less than X . A more unequal distribution would minimize the number of persons receiving an allotment of less than X , but would ensure that the least advantaged would have less than under the difference principle, and would lead more miserable lives. Assume that enough of this information is known to the parties to allow them to infer that each is more likely to receive an allotment of X or greater if the more unequal distribution is chosen instead of the Rawlsian one, but the information is not complete enough to allow them to infer how much more likely. Here, following a maximin strategy would not appear to be rational. It seems to be worth risking receiving a very low level of basic primary goods, in order to obtain a better chance of receiving an allotment that would enable one to live a life that is minimally decent and worth living. To follow MMP, under these circumstances, would be to maximize one's chances of obtaining a disastrous and unacceptable outcome, and this does not seem reasonable. Hence, the propriety of applying MMP in the Rawlsian choice situation is decidedly less plausible when that situation is revised to resemble the deter-or-disarm choice situation in relevant respects. This suggests we should look beyond (modified) MMP in seeking a solution to our problem.

III. DISASTER AVOIDANCE

No generally satisfactory principle of rational choice under uncertainty has yet been found.¹⁵ Fortunately, for our purposes, we only need a principle that is plausibly applicable to a very limited class of cases, those having the same relevant structural features as the deter-or-disarm choice. To aid discovery of such a principle, let us consider an example involving rational prudence.

A forty-year-old man is diagnosed as having a rare disease and consults the world's leading expert on the disease. He is informed that the disease is almost certainly not fatal but often causes serious

paralysis that leaves its victims bedridden for life. (In the other cases, it has no lasting effects.) The disease is so rare that the expert can offer only a vague estimate of the probability of paralysis: 20 to 60 percent. There is an experimental drug that, if administered now, would almost certainly cure the disease. However, it kills a significant but not accurately known percentage of those who take it. The expert guesses that the probability of the drug being fatal is less than 20 percent, and the patient thus assumes that he is definitely less likely to die if he takes the drug than he is to be paralyzed if he lets the disease run its course. The patient would regard bedridden life as preferable to death, but he considers both outcomes as totally disastrous compared to continuing his life in good health. Should he take the drug?

The choice facing this patient has the same basic structure as the deter-or-disarm choice and is essentially like the choice facing the parties in the revised Rawlsian situation described above. If he ignores the very tiny chances of nontreatment resulting in death and treatment being followed by paralysis,¹⁶ he must choose, under two-dimensional uncertainty, between a smaller probability of a larger disaster (the chance of death after taking the drug, which would correspond to large-scale nuclear war after deterrence) and a larger probability of a smaller disaster (the chance of paralysis after nontreatment, which would correspond to Soviet world domination after U.S. nuclear disarmament). Let us imagine the advice he would receive from three different friends (who also ignore the aforementioned "tiny chances"): an expected utility maximizer, a maximiner, and a disaster avoider.

The expected utility maximizer's advice: To choose rationally is to maximize expected utility. In this case, carefully consider the outcomes – death, bedridden life, and normal life – and estimate their relative values to you. Then estimate as best you can the probabilities that you will die if you take the drug and be paralyzed if you do not. Using these estimates, calculate the expected utility of each course of action and choose the one with the higher expected utility. Admittedly, it would be better if more information on the disease and the drug were available, but nonetheless, following EUP is the best you can do.

The maximiner's advice: Since you must choose essentially under conditions of uncertainty, you should act to make sure the worst does not happen. It would be silly to risk death for an indefinite

improvement in your chances of continuing a normal life. At least if you refuse the drug, you know that the worst that might happen to you is paralysis.

The disaster avoider's advice: Since you regard paralysis as an extreme personal disaster, it would be wrong to sacrifice the extra chance of continuing a normal life, even to avoid the risk of death. Admittedly, if this were one of a series of like choices, the matter might be different. For you would quite likely perish if you were to take (different) potentially fatal drugs in each of a series of like cases. Since, however, this is a once-in-a-lifetime choice, and there is a very good chance of success, you ought to give yourself the best chance of obtaining an acceptable result and take the drug.

The first two friends advise, respectively, that the patient follow EUP and MMP. The third friend's advice conforms to the following Disaster Avoidance Principle (DAP): *when choosing between potential disasters under two-dimensional uncertainty, it is rational to select the alternative that minimizes the probability of disaster occurrence.* The reasoning of the third friend seems at least as cogent, and his advice seems at least as attractive, as the reasoning and advice of the expected utility maximizer and the maximiner. This suggests that DAP ought to be taken seriously, as an alternative to EUP and MMP, as a principle governing choices between potential disasters under two-dimensional uncertainty.¹⁷ Further, it appears that DAP will have greatest credibility (compared to EUP and MMP) when all of nine special conditions are satisfied.¹⁸ Each of these conditions is listed below (in italics), followed by a brief explanation of why it lends relative plausibility to the application of DAP.¹⁹ (To preclude any possible misunderstanding, it should be emphasized that the following statements are not axioms from which DAP may be derived, but rather are statements of the limiting conditions for highly plausible applications of DAP.)

- (1) *The chooser lacks reliable quantitative probability and utility estimates.* Under this condition, principles of choice such as EUP that rely on such estimates are relatively less attractive.
- (2) *The chooser has confidence in his ordering of the (conditional) probabilities of the various outcomes.* This suggests that any principle such as MMP that ignores ordinal probability data is relatively less attractive.
- (3) *All disastrous outcomes are regarded as extremely unacceptable, that is,*

as involving very large amounts of net negative utility. The truly disastrous nature of the lesser disaster makes it sensible to risk the worse disaster in hope of avoiding all disasters. The wisdom of following MMP is quite doubtful under these conditions, as this would maximize the probability of there being an extremely unacceptable outcome.

- (4) *The disastrous outcomes are judged to be of roughly the same order of magnitude, that is the worse disaster may be many times worse than the lesser disaster, but it is not hundreds of times worse (or more).* If the larger disaster were a hundred or a thousand times as bad as the lesser one, principles such as MMP that emphasize disaster minimization rather than disaster avoidance would seem more attractive.
- (5) *The chooser regards the utility disparity between the nondisastrous outcomes (e.g., between the status quo and mutual disarmament) as being small compared to the utility difference between the disastrous and non-disastrous outcomes.* If this were not so, a principle such as DAP that ignores the relative desirability of the different nondisastrous outcomes would be less attractive.
- (6) *The choice is unique, that is, is not one in a series of like choices.* For a series of like choice, averaging principles like EUP are relatively more attractive
- (7) *The probabilities of the disasters are not thought to be insignificant.* If the risks of disaster were thought to be negligibly small, principles such as EUP that take into account the relative merits of the nondisastrous outcomes would be relatively more attractive.
- (8) *The probability of the greater disaster is not thought to be very large.* When this probability seems quite high, disaster avoidance is a rather forlorn hope and its importance pales in comparison with the goal of disaster minimization.
- (9) *The probabilities of the disasters are not thought to be very close or equal.* If they were thought to be very close, following a strategy of disaster avoidance rather than disaster minimization would seem less plausible.

Are these nine conditions satisfied (or very nearly satisfied) by the deter-or-disarm choice situation? I believe it is not unreasonable to suppose that they are, though obviously the applicability of many of them could be debated at length. Rather than enter into such a

debate, I shall simply state what each of the conditions would amount to when applied to the deter-or-disarm choice, as viewed from the perspective of the United States, and make a few brief observations about some of the more controversial conditions. The reader may then judge for herself the plausibility of regarding the deter-or-disarm choice as satisfying these conditions.

First, we do not have reliable estimates of the utilities of large-scale nuclear war and Soviet world domination, or of the probabilities that these outcomes would result from U.S. minimum deterrence (for thirty years) or U.S. nuclear disarmament. Second, we can be confident that the likelihood of Soviet domination if the United States disarms is greater than the likelihood of war if the United States practices minimum deterrence. Third, both war and Soviet domination would produce extremely large amounts of negative utility. Fourth, Soviet domination would produce negative utility of roughly the same order of magnitude (i.e., not hundreds of times less) as would war. Fifth, the amount of positive utility produced by U.S. nuclear disarmament, if it did not lead to Soviet domination, would be small compared to the amount of negative utility produced by either Soviet domination or war. Sixth, since (i) either disarmament or deterrence could lead to a disaster that would eliminate the opportunity to make similar choices in the future, and (ii) the circumstances surrounding the balance of terror will likely be different in thirty years in any case, the present deter-or-disarm choice should not be treated as simply one in a series of like choices.²⁰ Seventh, the probabilities of U.S. nuclear disarmament leading to Soviet domination and U.S. minimum deterrence leading to war are not so small as to be disregarded. Eighth, the probability of U.S. minimum deterrence leading to war is not very large. Ninth, the probabilities of U.S. nuclear disarmament leading to Soviet domination and U.S. minimum deterrence leading to war are not very close to equal.

Some reasons for believing the second and ninth statements are given in Chapter 6. The eighth statement is highly plausible because superpower leaders appreciate the awful consequences of nuclear war and because the main risks of nuclear war under current policies would be largely avoided by a policy of minimum deterrence.²¹ The other statements are minimally controversial, save for the fourth which claims that large-scale nuclear war would not be hundreds of times worse than Soviet world domination. In evaluating this state-

ment, we should note that such domination (a) might itself involve significant nuclear attacks, (b) could be expected to be long-lasting in view of modern methods of gathering and controlling information, and (c) would likely be replaced by other authoritarian systems if and when it broke down. Also, while large-scale nuclear war would be an enormous calamity for humankind, it seems on current evidence that it would not be likely to lead to human extinction,²² and humankind might well recover from most of the damaging effects of such a war in the very long run. In light of these considerations, it may be reasonable to accept the controversial fourth statement.

If the nine statements are true, or very nearly so, then the deter-or-disarm choice (faced by the United States) is one to which DAP can be applied with considerable plausibility. Such application yields the conclusion that, from a utilitarian point of view, minimum deterrence is more rational than unilateral disarmament, and hence is morally permissible.²³ To lend support to this conclusion, certain objections to DAP and its application that might be offered by advocates of EUP or MMP are considered in the next section.

IV. OBJECTIONS

A plausible minimum requirement for a principle of choice being satisfactory is that it guarantees transitivity.²⁴ One might contend that, because the concept "roughly the same order of magnitude" used in the fourth condition of application of DAP is intransitive, DAP will not satisfy this requirement. Consider a choice between disaster-risking acts *A*, *B*, and *C*, in which the other eight conditions are satisfied, and the probabilities and magnitudes of the disasters that may follow from each of the three acts rank in reverse order (with *A* being least likely and *C* most likely to lead to disaster). Suppose *B* risks a disaster that is of roughly the same order of magnitude as the disasters risked by *A* and by *C*, without the latter two disasters being of roughly the same order of magnitude. It may appear that transitivity breaks down in such a case. For while DAP ranks *A* above *B*, and *B* above *C*, it appears to stand mute on the comparison between *A* and *C*, because the fourth condition of its application fails to hold.

In response, it should first be noted that DAP itself is transitive and would imply that *A* is to be preferred to *C*. The transitivity

problem arises only with respect to applications of DAP, when we restrict the conditions for such applications in order to ensure high plausibility. Further, this problem does not appear to be unique to DAP, and may be a general feature of ordinal principles of choice under uncertainty. MMP cannot plausibly be applied to all choices under uncertainty, and certain sensible restrictions on its application would lead to a similar breakdown of the guarantee of transitivity. (This would occur, e.g., if we refused to apply MMP to choices between two acts, whenever the act with the worse worst outcome has a median value – that is, a value halfway between its best and worst outcomes – that is *n* times larger than the median value of the act with the better worst outcome.)

While this observation might inhibit use of the transitivity objection by maximinners, supporters of EUP may conclude that it shows that ordinal principles in general are inadequate principles of choice under uncertainty. Such persons may be more tolerant of another way of dealing with DAP's transitivity problem: viewing DAP, under the given conditions of application, as an approximation of a more complex nonordinal principle of rational choice. My candidate for such a principle is a sort of weighted average of DAP and EUP that may be called the Compromise Principle (CP).²⁵ Suppose one must choose between various acts under conditions of two-dimensional uncertainty. One estimates as best one can the utilities and probabilities of the various outcomes, then regiments the utility scale so that the utilities of the outcomes vary between zero and one.²⁶ Let EU_j be the expected utility of the *j*th act based on these probability and (suitably regimented) utility estimates, *r* be a measure ranging between zero and one that represents one's level of confidence in one's probability and utility estimates, and PDA_j be the probability that the *j*th act will *not* result in disaster. To each act *A_j*, assign an index $C_j = r(EU_j) + (1 - r)(PDA_j)$. The Compromise Principle says to perform the act with the highest index. When *r* = 0, we have no confidence in our quantitative estimates and CP is equivalent to DAP. As *r* increases from zero to one, CP diverges from DAP and more closely approximates EUP. When *r* = 1, we have a choice under risk, and CP is equivalent to EUP.

DAP may be viewed as an ordinal approximation of the nonordinal and transitive CP, in that the two are likely to yield the same choice in cases in which the special conditions for plausible application of DAP are satisfied. The satisfaction of condition one

implies that r , the confidence index, is quite low. Conditions four and five imply that the *EUF*'s are not likely to be too far apart, while condition nine says that the *PDA*'s are not very close together. Under these conditions, the second term of the index is likely to dominate the first, and CP and DAP are likely to yield the same prescriptions.²⁷

In summary, the transitivity objection does not seem to be a telling objection to DAP. DAP is transitive (though not necessarily always plausible) when applied across the board to all choices under two-dimensional uncertainty. While transitivity is not guaranteed when application of DAP is suitably restricted, this appears to be a general problem with ordinal principles of choice and not a characteristic flaw of DAP in particular. Further, it is possible to view DAP as an approximation of a nonordinal principle of choice that guarantees transitivity. Thus, DAP may plausibly be applied with assurance of transitivity when the nine conditions hold between the entire set of alternative acts. If they hold between some pairs of alternatives and not others, CP can be applied to transitively rank the alternatives.

Introduction of CP aids us also in dealing with another objection that might be voiced by expected utility maximizers: that DAP is plausible only because its application has been limited to cases in which it agrees with EUP. Now it is true that some of the conditions of application were introduced to ensure that DAP is not applied when there is too massive a divergence between it and EUP. But the two principles will not necessarily agree, even when the nine conditions are satisfied. Viewing DAP as an ordinal approximation to CP makes clear that the main idea behind DAP is to diverge from EUP, by hedging against disaster occurrence to the extent required by the unreliability of expected utility calculations. Thus DAP is not simply a disguised version of EUP. Supporters of EUP may feel that this amounts to an admission that the use of DAP is irrational, but I do not see why this should be so. The disaster avoidance approach embodied in DAP represents an attempt to deal with some choices under two-dimensional uncertainty rationally, from a utilitarian perspective, while avoiding both of two opposite mistakes: using quantitative methods without the necessary quantities, and ignoring utility altogether because of our inability to apply precise quantitative methods.

Consider, finally, an objection that a maximiner might offer:

"Generally, one has an obligation to act more conservatively with respect to imposing risks on others than with respect to taking risks oneself. Therefore, DAP has considerably less plausibility as a utilitarian principle of choice regarding imposing risks on others, than it does as a principle of rational prudence."²⁸ In reply, it may be noted that conservatism, as used in the principle cited in this objection, generally means an aversion to risking losses in hope of obtaining uncertain gains. DAP applies, though, only in situations in which there is no sure way to avoid the risk of disastrous losses, and the only question is whether to minimize the degree or the probability of such losses. In such cases the usual notion of conservatism does not apply. Once this is seen to be so, it becomes apparent that in choosing between potential disasters, it shouldn't matter whether the potential victim is oneself or another. Imagine, for example, that the patient described in section III is your ward and is unconscious. You must decide whether she is to receive the drug, knowing that while she prefers paralysis to death, she regards either as an unmitigated disaster compared to retaining her health. The rational promotion of your ward's interests in this case should not involve appeal to principles that are different from those you would use to decide if you were the patient deciding prudentially. If it is rational for you to take the drug if you are the patient, it is moral and rational for you to ask that the drug be given to your ward when she is the patient. Hence, the maximiner's objection is answerable.

V. COMPLICATIONS

Our basic analysis considered only two outcomes for each alternative. At first, it may seem this biases the analysis in favor of deterrence. For it means, in effect, ignoring the possibility that one's rival's domination of the world following one's unilateral disarmament might be incomplete, or that substantial change for the better in the rival's system might accompany his domination. Further, no account is taken of the fact that a deterrent policy could lead – without war – to one's rival being dominant, as a result of his decisively winning the arms race or the accompanying ideological-political struggle. These features of the choice problem are not reflected in the above analysis, and they place the disarmament alternative in a more favorable light compared with the deterrence alternative. However, consideration of a wider range of outcomes

also reveals features of the choice situation favorable to the deterrence alternative. For, like domination of the world by one's rival, nuclear war admits of degrees: it would not necessarily be fought to the limits. Also, unilateral disarmament would no more constitute a guarantee against nuclear war than deterrence constitutes a guarantee against one's rival dominating the world. (Some, in fact, would consider unilateral disarmament a virtual invitation to nuclear attack by one's rival, an invitation that might be accepted.) Hence, it does not appear that simplifying our problem by considering only two outcomes for each alternative policy introduces a net bias in favor of a deterrence policy.

Another simplifying assumption – ignoring the existence of other nuclear powers – results in a substantially lower estimate of the likelihood of unilateral disarmament leading to nuclear war. In fact, once such third powers are taken into account it might be supposed that unilateral nuclear disarmament by one superpower (say the United States) would, by eradicating any semblance of strategic balance, produce a much greater risk of nuclear war (in some form or another) than would a policy of minimum deterrence. Hence, if anything, this assumption makes the deterrence option seem less attractive than it really is.

These last two observations may be brought together and amplified by considering an objection to our argument offered by Jefferson McMahan.²⁹ Once we acknowledge that both minimum deterrence and unilateral nuclear disarmament could produce either disastrous outcome (as well as the status quo), we cannot infer that the former policy is more likely to avoid disaster simply from the fact that its most likely disastrous outcome is less likely than the most likely disastrous outcomes of the latter policy. We need instead a comparison of the likelihood that minimum deterrence would produce *some disastrous outcome* (either large-scale nuclear war or domination) with the likelihood that unilateral nuclear disarmament would produce one of these outcomes. This objection is correct, but can be readily answered in either of two ways.

First, it is asserted in the ninth statement in section III (and argued in Chapter 6) that the probabilities of the single most likely disasters following from minimum deterrence and unilateral nuclear disarmament are *not very close*. At the same time, it is highly plausible to suppose that the probabilities of the less likely disasters occurring (i.e., of deterrence leading to domination or of unilateral disarmament

leading to large-scale nuclear war) are quite small. Thus, even if the former were larger than the latter, the difference between them would be smaller than the substantial difference between the probabilities of the most likely disasters. It follows that a minimum deterrence policy is more likely to avoid disaster than is unilateral nuclear disarmament. Of course, this argument employs “semi-cardinal” judgments in describing the probabilities of the less likely disasters as “quite small” and the gap between the probabilities of the larger disasters as “substantial.” But these judgments are both plausible and much weaker than the more precise cardinal assumptions needed to calculate expected utilities.

Second, using nothing but purely ordinal probability comparisons, we can reach the same conclusion about the superiority of minimum deterrence for disaster avoidance. This requires making the assumption that unilateral nuclear disarmament is more likely to lead to large-scale nuclear war than minimum deterrence is to lead to domination. Given, on the one hand, the likelihood that third nuclear powers (especially China) would remain armed (and increase their stockpiles) if the United States disarmed unilaterally, and, on the other hand, the grave dangers of (and lack of incentive for) attacking a minimum-detering United States, this assumption seems highly plausible. If it is true, purely ordinal probability considerations favor minimum deterrence according to DAP. For then minimum deterrence is less likely (than unilateral nuclear disarmament) to produce both its more and less likely disastrous outcomes, and hence less likely to produce some disastrous outcome.

Our last simplifying assumption is the setting of a thirty-year period for the operation of the chosen policy. Does this bias the analysis in favor of deterrence, by making the risks of war seem less than they really are? Admittedly, making the policy period very short might introduce such a bias. Thus if we choose *one day* as the policy period, and argued on Monday that we should practice deterrence because the risks of this one-day use of the policy are negligible, then argued on Tuesday that we should continue deterrence for another day for the same reason, and so on, we would be engaging in sophistry. However, the period used in our analysis is not so short that the risk of war resulting from practicing deterrence throughout that period is negligible, and our analysis does not rely on ignoring this risk, but rather on recognizing it and balancing it against the risks of the alternative policy.

Even if the thirty-year policy period is long enough, wouldn't repeated applications of the analysis at the beginning of each period lead to the recommendation to continue deterrence indefinitely? If so, perhaps the risks of continuing the policy in the very long run (e.g., for hundreds of years) should be taken explicitly into account. To suppose, however, that the same analysis can be applied at the start of each future thirty-year period is a mistake. Changing conditions may render deterrence either unnecessary or too dangerous, so that it can no longer be justified, in accordance with DAP, as the lesser utilitarian evil. Further, the overall argument of this book suggests that each side should specifically aim at bringing about changes of the former sort by seeking bilateral disarmament. In essence, then, our analysis is not intended to apply to the very long run, or to justify continuing deterrence regardless of how conditions change. Rather, it seeks to answer the moral question that seems most appropriate in the present historical context: given its perceptions of the political-military situation, is it permissible for a superpower to practice minimum deterrence for a significant period of time, if it uses this time to attempt to alter the conditions that (seem to) make deterrence necessary? No bias is introduced into the analysis by setting a definite time horizon for the purpose of answering this limited question.

We may conclude that the central implication of our analysis – the permissibility of minimum deterrence – is not a product of the simplifying assumptions introduced to render the basic structure of the deter-or-disarm choice clear. It is rather, I think, a conclusion that follows plausibly from a utilitarian thinking about the central elements of the deterrence problem.

A final observation is in order. One view of the balance of terror is that it results from each side selfishly pursuing its national interests, rather than adopting a moral posture and seeking to promote the interests of humanity as a whole. However, the arguments of this chapter indicate that given each side's perceptions of the present political-military situation, rational promotion of the general interests of humanity would recommend that each side practice minimum deterrence, at least in the short run. If this is correct, a rather surprising conclusion follows. Even a miraculous conversion to a general humanitarian morality by the United States and the Soviet Union (or their governments) would not, in itself, suffice to liquidate the world's nuclear danger. Rather, a solution to the

balance of terror must be achieved by a different process, one often advocated by conflict theorists: changing U.S. and Soviet perceptions of each other and gradually building mutual trust between the two nations and their governments, by means of a stage-by-stage bilateral nuclear disarmament process. In Chapter 10 it is argued that no structural features of their nuclear competition would prevent the superpowers from carrying out such a process, were they to act in a rational and farsighted manner.