



Nuclear Intentions

Author(s): Gerald Dworkin

Reviewed work(s):

Source: *Ethics*, Vol. 95, No. 3, Special Issue: Symposium on Ethics and Nuclear Deterrence (Apr., 1985), pp. 445-460

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/2381031>

Accessed: 02/11/2012 16:32

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*.

Nuclear Intentions*

Gerald Dworkin

A terrorist engages in violence in violation of law against people who do not understand themselves to be at war. The victims of terrorist attacks are unarmed, undefended, and unwary. The crucial point is that they conceive of themselves as civilians. They do not understand that they are regarded by the terrorist as belligerents in an ongoing war. Terrorist war is part of a total war, which sees the whole of society as the enemy and all the members of society as appropriate objects for violence.¹

This is an essay on the relationship between the morality of the use of nuclear weapons and the morality of the threat to use such weapons under certain conditions, that is, the doctrine of nuclear deterrence. There are four possible positions that can be taken about the permissibility of use or the threat to use. One can claim that both the actual use and the threat to use are morally permissible. This is the view of Gauthier.² One can claim that although actual use is wrong the threat to use can be morally legitimate under certain conditions. This is the view of Kavka.³ One can believe that neither the actual use of nor the threat to use nuclear weapons is morally permissible. This is the view I shall be defending in this essay. Finally, there is the view, held by nobody to my knowledge, that, although the actual use of nuclear weapons is morally legitimate, the threat to use is illegitimate.

In addition to the various substantive differences just mentioned there are differences in the logical relationships different philosophers use in their arguments. There are those, such as Kenny, who argue that because the actual use of nuclear weapons is impermissible so is the

* I wish to express my appreciation for comments on earlier drafts by Charles Chastain, John Deigh, Art Flemming, Robert Goodin, Russell Hardin, Chris Korsgaard, Richard Kraut, and Jeff McMahan. Versions of this article were read at the Arms Control Seminar at the University of Chicago, April 8, 1984, the Aspen Conference on the Ethics of Nuclear Deterrence and Disarmament, September 2, 1984, Aspen, Colorado, and the University of Colorado.

1. Jeanne J. Kirkpatrick, *Harper's* (October 1984), p. 44.

2. David Gauthier, "Deterrence, Maximization, and Rationality," *Ethics* 94 (1984): 474–96.

3. Gregory S. Kavka, "Some Paradoxes of Deterrence," in *Moral Issues*, ed. Jan Narveson (Oxford: Oxford University Press, 1983), pp. 72–87, p. 76.

Ethics 95 (April 1985): 445–460

© 1985 by The University of Chicago. All rights reserved. 0014-1704/85/9503-0005\$01.00

threat to use.⁴ There are those, such as Kavka and Lewis, who believe that the questions of use and threat are logically independent of one another.⁵ And finally there are those such as Gauthier and I who focus on the permissibility of threatening and argue that if the threat is legitimate (rational for Gauthier) then one does not need a separate justification for carrying out the threat.

For some parts of my argument I shall be assuming that the actual use of nuclear weapons is morally and rationally forbidden. While this is, as I have already indicated, not a position which enjoys unanimous support, it is one which can be plausibly supported by a number of different considerations and by quite different normative positions.

Under the traditional just-war criteria the use of weapons which cannot be used in a manner so as to discriminate between combatants and noncombatants is forbidden. One can quibble about the exact definition of "noncombatant," but in any reasonable view the millions of children under the age of fifteen, the millions of women engaged in running households, the millions of persons in hospitals and nursing homes, are not combatants. Even if one has a very broad notion of collective responsibility so that any citizen of an aggressor nation is somehow implicated in the aggressive action of its leaders, the citizens of other countries who would be damaged by radioactive fallout or, if recent predictions are correct, by the consequences of nuclear winter would have to be considered noncombatants on even the broadest view of the distinction.

Those who disagree with my assumption may regard my argument as hypothetical in form. If the actual use of nuclear weapons is neither rational nor morally permissible, what follows about the morality of the possession, threat, or conditional intent to use such weapons?

I

Under a policy of nuclear deterrence we do many things now, for example, appropriate money for missiles, manufacture them, install them, maintain them, train personnel to fire them, make announcements about them, and negotiate about their dismantling. In addition to all these acts and policies those in positions of responsibility for making policy and implementing it form intentions for the future. We do all the above in the light of a decision to act in certain ways, under certain conditions. We can, of course, change our minds about what we are prepared to do. We can give up our intentions for the future. Or we can keep them, and the world may go in such a way that they are never carried out, either because the conditions for their implementation do not arise or because we cannot do what we set out to do. But to adopt or form an intention for the

4. Anthony Kenny, "Better Dead than Red," in *Objections to Nuclear Defense*, ed. Nigel Blake and Kay Pole (London: Routledge & Kegan Paul, 1984), pp. 12–27.

5. Kavka; D. Lewis, "Devil's Bargains and the Real World," in *The Security Gamble: Deterrence Dilemmas in the Nuclear Age*, ed. D. MacLean (Totowa, N.J.: Rowman & Allanheld, 1984).

future is to commit oneself in certain ways—to action, to belief, to values and purposes.

Future intentions are objects of moral evaluation and assessment. Insofar as the formation or adoption of an intention is an act of the agent, although indeed a mental act, it is subject to evaluation as an act. One's decisions, plans, resolutions, have effects both on oneself and on others and can be assessed morally in terms of those effects. In addition, just as certain overt acts are wrong because of the kind of act they are, so can the formation of an intention be wrong because of the kind of intention it is.

Future intentions can also be evaluated as evidence for, or as constitutive of, the character of the agent. What an agent decides to do, what she is prepared to do, what she is willing to do, tell us what kind of person she is. And insofar as future intentions result in future acts (with the appropriate causal connection) we assign responsibility in terms of those intentions.

II

The question, in its broadest form, may be put this way. What is wrong with (conditionally) intending evil? What is wrong with intending to do that which it is morally forbidden to do? It seems that the answer to this question cannot be independent of one's moral theory. For the theory specifies the features in terms of which something can be right or wrong. For example, in a divine-command ethical theory at least part of what is wrong in intending to do what is divinely forbidden is that one is disobedient to the will of God. For consequentialist moral theories what makes it wrong to intend to do what is forbidden is that such an intention makes it more likely that one will act in a manner which will be productive of the bad consequences that are forbidden. For the Kantian, to intend that which is forbidden is to be willing to act in a manner which is ruled out by the Categorical Imperative. The intent is wrong quite independent of its being linked to future effects. For most deontological systems there will be a special importance attached to what we do intentionally as opposed to what happens as a result of what we do. The choice of evil becomes the focus of moral appraisal.

I do not intend to settle matters of fundamental moral theory here. I shall not argue what the relative importance of consequentialist versus deontological considerations ought to be. Instead I want to look at various kinds of deterrent strategies and explore how consequentialist or deontological theories would assess the role and significance of a conditional intention (or its absence) to use nuclear weapons on the assumption that their use is wrong.

III

I am going to consider six possible options. They have been selected because of their usefulness for exploring the moral issues and not because

they are the ones which strategists regard as the most realistic or practical options available to us at this point in time. Indeed, two of them are not technologically feasible currently.

1. *Deterrence*.—In this option the United States possesses a second strike capability and forms the conditional intention to use this capability to retaliate for any nuclear attack on the United States. It either announces this policy or makes the policy manifest in some other way. In either case it threatens an aggressor with retaliation under certain conditions.

2. *Bluff*.—This option is the same as deterrence except that we make a secret decision not to use the weapons under conceivable circumstances. We do not form the conditional intention to retaliate.

3. *Retaliation machine*.—Herman Kahn originated the idea of a Doomsday machine which was programmed so that if nuclear weapons were exploded over the United States a device would be automatically triggered that would destroy the earth. For our purposes suppose, instead of destroying all human life, it launches our missiles at the attacker's homeland.

4. *Mere possession*.—The United States would maintain exactly the same nuclear arsenal and delivery system as in deterrence but would form no intention one way or the other as to future use. For this option to be perceived as different from 1 or 2 it would have to be made clear, in some fashion, that we are neither bluffing nor threatening. We merely possess a capacity without any plans or decisions about future use. By this I mean not that we make it unclear whether we have plans but that we make it clear that we have no definite plans.

5. *Autoretaliation*.—This terminology is borrowed from Kavka, as is the following description of the device: "Imagine that the U.S. invents a radio device that fifty per cent of the time is able to deflect Soviet ballistic missiles in flight and send them to present targets. For purposes of deterrence, the U.S. programs Soviet cities instead of the oceans as targets, and announces this openly."⁶ The point of the 50 percent limitation is that this is not a shield by which we can protect ourselves from harm but a way of inflicting harm on an aggressor.

6. *Bounce back*.—This is a form of shield which bounces back incoming missiles to wherever they are fired from. We do not have the capacity to target the deflected missiles. They simply return to wherever they are fired from. If they are fired from centers of population, they return to those centers. If they are fired from remote military bases, they return to those bases.⁷ Let us suppose that we know that, in fact, many of the enemy missiles are located

6. Gregory S. Kavka, "Nuclear Deterrence: Some Moral Perplexities," in MacLean, ed.

7. Douglas Lackey called my attention to the existence of a similar device (in his "Nuclear Deterrence and the Deflection of Risks" [1984, typescript]). Richard Ullman suggested to me a variation of this device which would deflect missiles targeted at civilian centers back to civilian centers and missiles targeted at military sites back to military sites.

around centers of population. For purposes of keeping things constant I assume that as in case 5 there is a 50 percent limitation on effectiveness.

These examples are selected because the existence and nature of various intentions varies among them. At the same time other variables such as risk to innocents, foreseen consequences, and the announcement of threats vary as well. By examining the moral evaluation of the different cases by differing theories we may hope to gain a better understanding of the role of intention.

IV

In case 1—the standard case of deterrence—we have the intention to retaliate and hence, by our assumption, the intention to do what is morally impermissible. The argument for the permissibility of the threat is a consequentialist one. The presence of the intention is assumed to be necessary for the credibility of the threat to retaliate, which is assumed to lower the probability of enemy attack. If that initial probability is high enough, if the losses due to such an attack are great enough, if the lowering of the probability is large enough, then it is plausible to suppose that the expected utility of forming the intention to retaliate is greater than the expected utility of some other policy which does not involve forming such a conditional intention. Of course such a utility calculation has to take into account all the likely consequences of such a policy, including the dangers of accidental war, the psychological effects of threatening behavior on one's adversaries, the effect of making such threats on those who make them and are prepared to carry them out, et cetera. And as the recent discussion by Hardin and Lackey has shown, there may be great disagreement about the results of such calculations.⁸ (Much such discussion has the appearance of somebody trying to figure out an answer to a mathematical question, but closer scrutiny almost always shows values being more or less arbitrarily assigned to both probabilities and utilities. As one State Department official put it in conversation about "real" defense calculations, "In all the years of talk about calculated risk I have never seen a calculation.") But the question I am interested in is what significance the presence or absence of conditional intentions plays in such a consequentialist assessment.

The answer must be, for such a theory, that the forming of the conditional intention to do x has causal consequences. Let us distinguish the consequences into the following classes. First, the probability of the agent doing x alters. For rational and noncompulsive agents what they wind up doing is some function of what they plan or intend to do. Other things being equal, it is more likely an agent will do what he intends to

8. Douglas Lackey, "Missiles and Morals: A Utilitarian Look at Nuclear Deterrence," *Philosophy & Public Affairs* 11 (1982): 189–231; Russell Hardin, "Unilateral versus Mutual Disarmament," *Philosophy & Public Affairs* 12 (1983): 236–54.

do than what he has no intention of doing. Second, the adoption of the conditional intention has what Kavka calls autonomous effects. It may, for example, make the subject more callous or more courageous. It may make others more fearful or more sanguine. Given that on consequentialist grounds the doing of x is ruled out, it would normally follow that an act which increases the probability of doing x is also ruled out. It is only because the act which increases the probability of doing x also increases the probability of other (good) consequences that the normal inference does not hold—or so Kavka argues. Since a consequentialist theory is concerned with the goodness or badness of states of affairs, the relevance of the forming of an intention is exhausted by its causal contribution to the production of one or another state of affairs. It makes no sense to assess the intention as being intrinsically wrong, although the willingness to adopt certain intentions may be good evidence about the moral character of an agent. In short, the relevance of intentions to do morally forbidden acts is exhausted by the increased risk of harmful consequences.

The difference between case 1 (deterrence) and case 2 (bluff) is exhausted, from the standpoint of evaluating intentions, by facts such as that under a bluffing policy there is no risk of doing the forbidden act, that there is a risk that the bluff may be exposed, that one will have to lie, and so forth.

Case 3 (retaliation machine) differs from 1 only in the fact that it is impossible to alter one's intentions. This, of course, affects the issue of risk in a complicated fashion. It decreases the risk of doing x since the threat to retaliate is made more credible, and this decreases the risk of attack. On the other hand it increases the risk of doing x since the option of changing one's mind, if attacked, is no longer present.

Lackey uses this case in an argument designed to show that Kavka is mistaken in supposing that it can ever be morally permissible to form the conditional intention to do x if x is itself morally impermissible: "To see that this is so, consider whether it would be permissible for the agent to construct a device such that, if C occurs, he cannot fail to do x . If it is not permissible to do x if C, it is not permissible to construct a device which will force him to do x if C if there is a real chance that C will occur."⁹

This obviously begs the question, for Kavka, if he is consistent, must believe that it is permissible, other things being equal, to construct such a machine. After all, he says that his agent is "willing, in order to prevent the offense, to accept a certain *risk* that in the end, he will apply the sanction."¹⁰ This must be the correct view for a consequentialist like Kavka since the machine merely fixes some level of risk that is antecedently decided to be acceptable.

9. Douglas Lackey, "The Intentions of Deterrence," in *Nuclear Weapons and the Future of Humanity: The Fundamental Questions*, ed. Steven Lee and Avner Cohen (Totowa, N.J.: Rowman & Allanheld, 1984).

10. Kavka, "Some Paradoxes of Deterrence," p. 76.

Case 4 (mere possession) presents a level of risk, relative to intention, which is intermediate between cases 1 and 2. It is lower than 1 insofar as it is left open that we will not retaliate. It is higher than 2 insofar as it is left open that we will. Ignoring autonomous effects, we create less risk that we will do the forbidden act than we do by forming the intention. Therefore it is a superior policy relative to deterrence.

Case 5 presents an interesting twist for the consequentialist since the causal effects on the enemy of adopting this policy are (or could be, depending on various technological assumptions) identical to those in case 1. In both cases we arrange matters so that missiles fall on the enemy, and in particular on noncombatants. The difference is in whose missiles do the falling. In the case of deterrence we launch our missiles at the enemy. In case 5 we deflect their missiles to our targets. In either case it is the launching of their missiles which triggers the retaliation. The example has the additional feature that our losses are (probabilistically) cut in half so that total damage is likely to be less than under a policy of deterrence, but this is not the feature I want to focus on.

Traditional theories of self-defense would distinguish between these policies. If someone throws a hand grenade into my house, I am entitled to throw it back at him, but am I entitled to throw it into his neighbor's house?

Case 6 is also an interesting one from a consequentialist standpoint. For in such a case we may have exactly the same degree of risk imposition as in case 1 or 5 but without forming the conditional intention to impose such risk. We do not announce that we will launch missiles or drop bombs on the aggressor or even redirect his missiles to his cities. We simply arrange matters such that the aggressor's own weapons fall on his own cities if he launches them from those cities. We cannot deny that we increase the risk of death and destruction to noncombatants, but we can deny that we have any intention of retaliating. We need not even possess offensive weapons. On consequentialist grounds the presence or absence of such an intention is irrelevant. Case 6 is not morally more favorable than case 1 from the standpoint of forming an intention. It may, of course, be more favorable on other grounds such as being cheaper, reducing the risk of accidental war, being psychologically less threatening, et cetera.

Case 6 provides us with a test case for assessing the difference between consequentialist and nonconsequentialist views about the moral relevance of intention since it separates out the elements of risk and intention which normally go together.

V

I now want to consider these various cases from a nonconsequentialist standpoint, that is, one which denies that the right action is always to be defined in terms of the production of the best consequences. An action may be wrong because of the kind of action it is, or because it cannot

be justified to those whom it adversely affects, or because it is not in accordance with a set of principles which would be consented to by all those affected by the action, or because it violates certain rights that individuals have. I leave unspecified the exact nature of the theory. For such views what is the significance of intending to do what it is wrong to do?

Roughly speaking, the significance is in terms of how the interests and welfare of other moral agents enter into our practical reasoning about what to do. Our values and our character are shown not by what comes about as a causal result of our actions but by the forms of reasoning we use in deciding how to act. Of course consequences count insofar as they must enter into our deliberations, but it is as intended or foreseen or negligently omitted from our deliberations that they enter. Let us look at cases 1 and 5 on the one hand and case 6 on the other from this standpoint.

It must be conceded that the installation of the bounce-back machine creates an increased risk of the same death and destruction to enemy noncombatants that a policy of deterrence does. If it is morally impermissible to impose such a risk intentionally, how can it be morally permissible to impose such a risk as a foreseen consequence of one's otherwise permissible acts? This, of course, is the issue usually discussed under the heading of double effect. I do not intend either to explicate or to defend that doctrine here. But I do intend to try and defend a view of the significance of intention which at least clears the way for defending the doctrine. For the initial conceptual obstacle seems to be the difficulty of seeing how the distinction between what one aims at as opposed to what one foresees can make any moral difference. There is an important difference between cases 1 and 5 and case 6 having to do with the form of practical reasoning we engage in when adopting one or another of the strategic policies. The difference between an agent who adopts a strategy of deterrence or autoretaliation on the one hand and one who adopts a strategy of a bounce back on the other is reflected in the means-ends relationships he is willing to countenance in order to attain his ends. In all cases I take the end to be that of preventing an unjust act of aggression, and I take that to be justified. In case 1 the agent imposes a risk of death and destruction on those who (by assumption) are not legitimate targets of such risk, and it is precisely by imposing such a risk that he believes his end will (causally) be obtained.

It, of course, does not follow that he desires to follow that causal path or that he would not prefer to obtain his ends in an alternative fashion. If, say, bluffing would work (almost) as well, he might prefer to bluff. But having made a determination that other causal paths would not work, he adopts this one. He uses threat and the imposition of risk to obtain his ends. In doing so he does more than show himself prepared or willing to risk the morally forbidden. He adopts a plan which makes the imposition and maintenance of such risk necessary (causally) to his being successful.

That this is so can be seen by considering the ways in which the agent must think and act in order to preserve the causal link between the unjust risk imposition and the end of successful deterrence. In both cases 1 and 5, if the potential aggressor were, say, to develop a successful antiballistic missile (ABM) defense, then we must seek to develop weapons which will defeat that defense. If the potential aggressor seeks to disperse the civilian population, then we must find ways of launching our own missiles or deflecting his so as to destroy those populations. The point is that, given our plan, the risks to the noncombatants must be maintained or increased. Their potential death and destruction is required for our plan to be successful. Otherwise our threats are empty.

And it is interesting to see that this is true of both cases 1 and 5. For both are based on the threat to inflict harm on noncombatants in order to deter. Whether there are other differences because of the fact that in one case it is our missiles that inflict the harm and in the other it is theirs I leave open. From the standpoint of intention the cases are parallel. In both we adopt a plan, and arrange causal relations, that requires the imposition of risk on those who (by hypothesis) are not the legitimate bearers of such a risk.

Whereas, in case 6, although we arrange matters in such a fashion that we may be able to foresee the same degree of injury, we do not require the imposition of such a risk in order for our plans to be successful.

If the enemy relocates his missiles far from population centers, we need not rearrange our strategy. Our end is not to be achieved through the imposition of risk on noncombatants, although we certainly foresee its occurrence. Of course if the enemy develops an ABM system, then, since our bounce back is only 50 percent effective, the absence of a deterrent system will leave us at a severe strategic disadvantage. But this is a separate point. All I am trying to show is that case 6 may be evaluated differently morally than case 5 and that the difference reflects what our intentions are.

I have claimed that the bounce-back case should be evaluated more favorably than the autoretaliator case and that it is permissible to employ it. There are, nevertheless, limits on what one can do even by way of simply deflecting threats. Here are two other cases for further reflection. First, suppose the bounce-back system deflected missiles onto the citizens of Switzerland. I am committed by my position to thinking this permissible since the noncombatants of the aggressor nation are no more responsible for the attack of their leaders than are the citizens of Switzerland. Second, suppose the incoming missile is targeted for a small city (80,000 population) but fired from a large one (8,000,000 population). Is it still permissible to deflect?

VI

Let us look at the differences between cases 1 and 2, deterrence and bluff. Jeff McMahan in his essay "Deterrence and Deontology" argues that the crucial moral element is the likelihood of using nuclear weapons,

not the presence of a conditional intention.¹¹ In his discussion of the bluffing case he argues that the deontologist is committed to being opposed to bluffing since those in the military chain of command would have to be deceived about the policy in order to maintain the illusion of deterrence. They would have the conditional intention to retaliate, although the commander in chief knows that he will in no circumstances issue an order to them to fire the weapons. But if it is morally impermissible to form the conditional intention to use the weapons, then it must be similarly impermissible to implant that intention in the minds of others. So the deontologist is committed to denying that there is any difference between cases 1 and 2. This is absurd since it condemns equally a policy in which there is no possibility that possessing nuclear weapons will lead to their use and one that involves such a risk.

This argument assumes a position which is far stronger than the one I am defending here. It assumes a deontologist who holds that all morally forbidden acts are equally wrong and a position which works only with the categories of the morally forbidden, the morally permissible, and the morally obligatory. It is more plausible to argue from a position which claims that two acts which are morally impermissible may differ in their wrongness and that our appraisals of persons and their characters are related in a much more complex way to the aforementioned division of acts.

If we are judging the commander in chief, then the morally crucial feature of his policy is that he is not prepared or willing or intending to bring injury to innocent parties. And his policy is not one that involves using potential risk to those parties in order to achieve his ends. It is true that he lies to his subordinates, but this shows only that it is not always morally forbidden to lie. And this must be assumed for any policy that involves bluffing.

As to those further down in the chain of command, it is not the case that they have the conditional intention to retaliate. They have the conditional intention to obey the orders of their superiors. This, indeed, involves the abdication of a certain amount of autonomy, but that is a feature of all military (and many other) structures. Whatever corruption is involved in renouncing the right to think for oneself about what one does is not associated in any special way with the bluffing scenario.

Again, in my view it may be impermissible to form the conditional intention to use nuclear weapons, and it may be impermissible actually to use such weapons, but I am not committed to the view that they represent equal wrongs. For why may not both be above some threshold which rules them both out and yet one be, so to speak, more above the threshold than the other? To say that they are both impermissible is to say that there is no adequate justification for either, or that neither would be allowed by any set of principles that meets our criteria for an adequate

11. Jeff McMahan, "Deterrence and Deontology," in this issue.

moral theory, or whatever. But that still allows that one of the acts is more evil than the other, in the sense that more bad is produced, and also allows for the view that if one had to choose one or the other, in a situation of necessity, one ought to choose the less evil. Of course if one has to choose between risking a nuclear war and fighting one, it is preferable to choose risking. If one had to choose between firing a pistol with six bullets in the chamber at a baby and firing a pistol with five bullets, one would choose the latter. But both are morally impermissible.

VII

If one believes that cases 2 and 6 are permissible, that case 4 is problematic to the extent that one leaves it open that one will retaliate, and that the other cases are not permissible, and if the alternative hypothesis of increased risk to the innocent does not explain all these judgments, it seems as if the focus remains on the conditional intention to retaliate. I want to set out some significant features of deterrent intentions which are morally relevant.

The first feature is that since the intention is conditional the intention creates a risk of doing what is morally forbidden, depending on the likelihood of the condition's being satisfied. The intention is only carried out if certain conditions are fulfilled. Actually this is probably true of most intentions which are thought of as "categorical." My intention to repay a loan is often conditional on the loan's not being forgiven. The doctor's intention to lie to his terminally ill patient is conditional on his being asked for the information. My intention to write a paper for a conference is conditional on my children's not becoming dangerously ill. Almost all intentions have an other-things-being-equal character to the commitment to fulfill them, even when this is not explicit.

Still it would be a mistake to think of all intentions as "really" being conditional in nature. This would be like thinking my belief that my pet is a cat assumes that I won't discover its insides to be made of transistors. In both cases there is a context within which the formation of the intention (belief) takes place, but the context is not part of the structure of formation.

That one has a conditional intention is, as is the case of nonconditional intentions, shown partly by how one's actions are guided by the intention—whether one takes appropriate steps to insure that one will be in a position to do what is envisioned if the conditioning event occurs, whether one takes steps to be in a position to determine that the conditioning event occurs, and so forth. With many conditional intentions the carrying out of the intention (if the conditioning event occurs) is desired for its own sake (If I get a raise, I will go to Paris for a week) or as a means to some end the agent has (If the test shows a malignancy, I will have an operation). The forming of a conditional intention shows what the agent is prepared to do and therefore what values guide his decisions.

The peculiar feature of deterrent intentions is that it is believed that the formation and expression of the intention makes it less likely that

the condition will occur. Since this is the point of forming the intention, it is part of the logic of deterrent intentions that one does not have to value the fulfillment of the intention, either as an end in itself or as a means to some other end one has. It is the forming of the conditional intention itself which reflects the agent's values by showing what he is prepared to do (under certain circumstances).

This last point is important and often misunderstood. For it is supposed that if deterrent intentions fail, that is, do not succeed in avoiding the occurrence of the condition, then the fulfillment of the intention must be viewed as a means to some other goal, the most usual one being future deterrence of the offender or others. But while this may be a reason for fulfilling the intention, it need not be. One may fulfill one's threat simply as a matter of keeping one's word or because the sanction deserves to be carried out.

Gauthier has argued recently that it can be rational to commit oneself to carrying out a threat, which apart from the context of strategic interaction would be irrational.¹² This view, which seems paradoxical to many, ought not to be. Schelling gives the example of the country which by making it clear that it intends to fight to the death makes it less likely that it will have to. And on a more mundane level any bargaining scenario demonstrates the same phenomenon in small scale. If I have something which I am willing to sell to you for \$100 or more, and you are willing to pay \$150 or less, there is room for mutual gain from trade. If these values are known to both of us, then if you offer me \$110 I may still commit myself to not selling for less than \$125 in the hopes that you will change your offer. If you do not, this may result in a situation in which, at that point, what is in my interest is to change my mind and accept your offer of \$110. But my best chance for getting the \$125 may be irrevocably to commit myself to not accepting your first offer.

It will not do to say that, when the stakes are so different, as in the nuclear context, the analogy cannot hold. For the difference in stakes only implies that it will be much more unlikely that the expected gains will outweigh the possible costs. But if they do, the same principles of rationality should apply.

What about the possibility of changing one's mind? Does this view imply that one cannot come to the conclusion that one has made a mistake? Of course not. But one must be careful about what one changes one's mind about. It is possible that one comes to the conclusion that one ought not to have formed the conditional intention in the first place, that it was not rational to be prepared to risk those costs for these benefits. But note that it cannot be the mere failure of one's strategy that one relies on to make this recalculation. For there was always the possibility of failure.

12. Gauthier.

Or, again, there is the possibility that, while then it was rational, given one's beliefs and values, to have made the commitment, one's values and beliefs have changed so that now one would not make the commitment.

Note also that one's opponent is entitled to take into account the likelihood of that kind of reassessment in calculating how likely it is that you will fulfill your commitment. He is also going to take into account how likely some form of nonrational backsliding is.

It also may be the case that one is not strong willed enough to be confident that one will carry out the threat in the knowledge that, at that point, there is no independent reason to carry it out. One then may seek to provide oneself with independent reasons (side bets) or to set up institutional mechanisms which insure the carrying out of the threat. But it seems to me an empirical question whether for a particular individual the existence of a good argument is by itself sufficient to provide him with enough motivation to carry out his intention or whether he needs external aids.

Gauthier goes on to argue the more controversial point not only that it is rational to form the conditional intention but also that the rationality of forming the intention shows that it is rational to carry out the threat if deterrence fails. For Gauthier there is no separate question that needs to be raised about the rationality of carrying out the intention since the rationality of the act is assessed as part of the policy of forming the intention *ab initio*.

I think it preferable to hold, as Kavka and Lewis argue, that, while the agent may not be irrational to carry out the threat since it stems from a commitment that it was rational to form, the act itself is irrational. By assumption it is an act which the agent has no independent reason to perform and reason not to perform. It is true that individuals who are able to make such commitments will achieve gains in utility that those who cannot must forfeit. But if the question arises whether one has reason to do what one threatened to do, given the failure of deterrence, the answer must be negative.

But the question I want to explore is whether the analogue for morality holds. Can it be moral to commit oneself to actions which, independent of the policy in which they are imbedded, are immoral?

In the case of deterrent intentions it is not enough to show that utility is maximized by the formation of such intentions, for that merely shows that it is rational to do so if one wants to maximize satisfaction of preferences. Nor is it enough to show that some important values, freedom or national security, are most efficiently secured by such intentions, for that merely shows that there is a moral reason in favor of such actions. But one function of morality is to put limits or constraints on what one may do to maximize satisfaction or to achieve important values.

If one takes as a heuristic notion here the idea of being able to justify one's actions to those whose interests are adversely affected by them, the

question becomes, Can one justify to those affected being prepared to take a risk of doing something the actual doing of which cannot be, independently, justified to them?

The central examples used by philosophers have been deterrent threats. The essential elements are that the coming about of C is up to those who are put at risk, they would be at fault if they brought about C, some important good is achieved by preventing C from being performed, and it is believed that the threat of x may prevent those put at risk from bringing about C. But there may be other kinds of cases which do not rely on deterrent aims.

Think, for example, of those who risk causing accidents by driving when drunk. Suppose we announce that such drivers will have to pay fines into a fund which goes toward the medical expenses of those who are injured in automobile accidents. We do this, not because this may have a deterrent effect (although it may), but in order to raise a fund to compensate accident victims. Since there is no causal connection between the behavior of those fined and any accident victim we are not entitled (in the absence of formulating and announcing such a policy) simply to confiscate property from such drivers on the normal tort grounds of making people bear the costs of their faulty conduct. But the charge of being used for a useful social purpose is, plausibly, defeated by the claim that it is not unfair to make those who voluntarily engage in faulty conduct pay the costs of those who are put at risk by their faulty conduct.

The crucial elements here are the fact that the carrying out of the intention is avoidable by the person put at risk, the fact that the condition which brings about the carrying out of the intention is not one which the person put at risk is entitled to bring about, and the fact that the person put at risk has been informed of the conditional intent.

A second case in which this kind of reasoning seems plausible is one in which one can address a justification to those whose interests are put at risk in terms of their interests having the best chance of being promoted by such measures. Suppose A is being threatened by B to stop conduct which is completely within A's rights. A is stubborn, will refuse to desist, and will suffer great harm from B. To avoid this I pretend to be so offended by the conduct that I threaten (sincerely) to punish A for the conduct in question. I believe he will refrain, not out of fear of my carrying out the threat (which I would carry out), but because he believes I am offended by the conduct and wishes to avoid hurting me. If, in fact, I am forced to carry out the threat (my policy fails), could I not justify this (otherwise immoral) action on the grounds that the policy of which it was a part provided the best chance to avoid harm to A?

These are essentially paternalistic cases where we impose a risk of bringing about (otherwise impermissible) harm as the best chance of avoiding worse harm.

The point in all these cases is that we start with a situation in which we are not entitled to do certain things to people. However, in these

situations we are entitled to threaten or warn that we will do these things unless those threatened or warned refrain from certain actions. The justification for the initial intention then carries over to the carrying out of the intention. We do not need additional moral reasons for being justified in carrying out the intention, although we may need practical reasons for being motivated to do so. Being entitled to perform some action does not necessarily give one a reason to do it. In the case we have been considering, the direction of moral justification is from the conditional intention to the carrying out of the intention. What needs to be established morally is that we are entitled to threaten or warn these agents in these situations.

It would seem then that there is the following parallel between rationality and morality. In both cases it can be justified to form a conditional intention, or to change one's dispositions, so that one is prepared to do what would otherwise be unjustifiable. Of course, in both cases there must be an argument to show that one is entitled to form the conditional intention in the first place. I turn now to the issue of deterrent intentions.

VIII

Since the intention in question is to inflict, under certain conditions, severe harm on other moral agents, there must be a justification which can be addressed to those agents as to why we are prepared to risk their harm. In the case where the person being threatened with harm is the person whose aggression against us we are trying to deter, the justification is obvious. Just as we would be entitled to use force in self-defense if we were in the process of being attacked, so we are entitled to threaten in advance in order to deter that attack. Whether the correct theoretical account of why we are entitled to use force in self-defense, and the derived entitlement to threaten to use force, is one in terms of the aggressor forfeiting his right not to be the object of force or the threat of force, or a notion of fault forfeits first, or the idea that we all stand to gain from an agreement permitting such force, is something I shall not attempt to settle here.¹³

It is also the case that there are certain limits which must be established on the amount of harm that potential aggressors may be threatened with. Again, I believe that these limits are established not by what we are entitled to do after the aggression but rather by what we would be entitled to do in the course of stopping the aggression, so that the justification for the threat explains the legitimacy of carrying out the threat rather than conversely. It is because it is impermissible to threaten noncombatants (on grounds of self-defense) that it is impermissible to bomb them, and not the other way round.

What about the other possible exception? Can we address to those

13. For some thoughts on this subject, see T. Hurka, "Rights and Capital Punishment," *Dialogue* 21 (1982): 647-60.

noncombatants who neither use force against us nor aid those who use such force, and whose welfare we risk by our deterrent intentions, an argument to the effect that they could have expected to gain from our policy of deterrence? The possibility of such an argument seems dubious both on empirical and on moral grounds. On empirical grounds we would have to establish that the expected gains to Soviet citizens (and citizens of other nations at risk of nuclear winter) from our policy of deterrence outweigh the expected risks to them. These gains cannot include a reduced risk of being exposed to our nuclear weapons, for a policy of unilateral nuclear disarmament assures them that that cannot happen. If it is argued that they have a reduced risk of being exposed to the nuclear weapons of other countries, for example, China, that assumes that we are prepared to threaten China with nuclear attack if they attack the Soviet Union. A policy which, as far as I know, we do not have.

The best argument along these lines is that the presence of nuclear weapons produces a greatly reduced risk of conventional war because of the fear of nuclear escalation. Since it doesn't matter a great deal to the average Soviet citizen whether he dies from a bullet or from a missile, if, in fact, the chances of being killed unjustly in conventional warfare (absent nuclear deterrence) were much greater than the chances of being killed unjustly in a nuclear war (with nuclear deterrence), there would be an argument along the lines I have suggested. I have no idea as to what the evidence for this assertion would be, nor do I think anybody else does.

But even if the facts were as indicated, it seems to me that there are a number of important moral objections to this argument. First, the facts are as they are only because it is assumed that we are willing to act immorally in the first place, that is, not to take much greater precautions than we do now to reduce the risks to noncombatants in conventional warfare. Second, such an argument only seems to me reasonable when it is addressed to members of one's own moral community. The argument is essentially paternalistic in form—we're doing this against your expressed will for your own (expected) good—and such an argument requires an ability to control, limit, and direct such powers by those whose liberty is being infringed.

Finally, and this brings us back to intentions once more, there is a special wrong where the risk is one of being the target of intentional harm. It is sometimes argued that as drivers we impose risks on pedestrians which are ultimately justified by the benefits that pedestrians gain by allowing the driving of automobiles. But I do not believe that we would accept an institution which imposed the same risks of injury and death that accidents cause, but risks brought about by actions aimed at the injury or death. Unlike an Aztec ritual of sacrifice or Shirley Jackson's lottery, we do not aim at the death of particular persons as the means of securing whatever benefits are at stake. In the case of deterrent intentions that is precisely what we do.