

Buy like a MADman, use like a NUT

When theoreticians think about nuclear deterrence, often they focus on a nasty choice between two rival package deals. The two have gone by various names over the years, but let me take the paired epithets: MAD (Mutual Assured Destruction) versus NUTS (Nuclear Use Theorists). Each package is a bundle of policies: centrally, policies for the procurement of strategic nuclear forces and conditional intentions about how to use those forces in case of war. I think we can break up the packages and keep only part of each. What we get may be in some sense MAD and in some sense NUTS – for the terms are elastic – but I hope it is the better half of both.

In a debate between MAD and NUTS, each side may say that the other's policies involve a twofold risk: a grave moral risk of committing massacres and a grave prudential risk of inviting and undergoing massacres. If they say so, they are right. The contest between these two repugnant alternatives gives nuclear deterrence itself a bad name. How does the very idea of nuclear deterrence turn into the nasty choice between MAD and NUTS? Does it have to happen? Is there no way around it?

First published in this form in QQ 6, No. 2 (1986), 5–8. This paper is an abridgement by Claudia Mills of David Lewis, "Finite Counterforce", in Henry Shue, ed., *Nuclear Deterrence and Moral Restraint* (Cambridge, Cambridge University Press, 1989), © Cambridge University Press 1989. Reprinted with kind permission from Claudia Mills and the Institute for Philosophy and Public Policy, and with the permission of Cambridge University Press.

MAD: IF YOU CAN'T BE CREDIBLE, BE DREADFUL

To trace the reasoning that drives us MAD, start with a simple conception of nuclear deterrence. We deter the enemy from doing X by threatening that if he does, then we will punish him by doing Y. But the enemy might notice that if he does X, we will then have no good reason to do Y. What's more, he may be able to give us a reason not to: he may threaten that if we do Y, then he will punish us by doing Z. Of course we may threaten that if he does Z then we will . . . but he might doubt that as well. In short we have a credibility problem: our deterrence is apt to fail because our threats are not believed.

How to solve it? One way is to make the threatened retaliation very, very severe. Then even if the enemy thinks we would have excellent reason not to retaliate, still he would not dare to call our bluff. If he evaluates risks as he should, multiplying the magnitude of the harm by the probability, we can make up in the first factor for what is lacking in the second. We can threaten a vast nuclear massacre, on an altogether different scale from the ordinary horrors of war. Destruction on this dreadful scale needn't be credible to deter. Although it could serve no good purpose to fulfill the threat, the risk that we might do so in blind anger suffices.

The MADman thinks it obvious that deterrence requires a solution to the credibility problem, and obvious that the only solution is to find a threat so dreadful that it needn't be credible; and he expects the enemy not to overlook the obvious. Therefore he thinks that for the enemy, as for us, an assured capacity to destroy cities will be seen as the *sine qua non* of nuclear deterrence. Further, he thinks it would take no great effort for the enemy to counteract any steps we might take to protect our cities. Therefore he thinks such steps would be, at best, costly and futile. We buy the means to reduce the enemy's strategic forces by counterforce warfare; the enemy buys enough more missiles to assure himself that enough of them will survive. We buy expensive defenses, the enemy buys enough more missiles to assure himself that enough of them will get through. We spend money, he spends money. Afterward there are many more nuclear weapons in the world, and each of them is one more place where an

accident could happen. In case of war, not only does the world get fallout and smoke from the destroyed cities, but also it gets fallout (and smoke and dust) from preliminary counterforce attacks and intercepted warheads. And still our cities are subject to vast and intolerable destruction. What have we gained?

The MADman boasts that his goals for deterrence are “finite.” If each side can count on having enough surviving weapons to meet the standard of assured destruction, that is all that either side has reason to want. Neither side has an incentive to expand or improve his forces, for all that would happen is that the balance would be reestablished at increased cost, increased risk, and increased danger to the rest of the world.

Thus the MADman’s policy for procurement of nuclear weapons is as moderate and benign as can be, short of renouncing nuclear deterrence altogether. But his policy for conduct of nuclear war is quite the opposite. What is the Commander-in-Chief supposed to do if deterrence fails? He is not supposed to do anything to protect the country entrusted to his care; he cannot, since it was thought futile to provide the means for limiting damage. Rather he is supposed to fulfill the threat to destroy cities – a vast massacre, serving no good purpose whatever. There is nothing else he can do. Thus MADness carries a grave moral risk. According to MADness, anything that can be seen to raise the chance of retaliation is all to the good. But it is all to the bad if deterrence fails: for what is raised is the chance of the most wicked act that it is possible for anyone in our time to perform.

NUTS: THE CREDIBLE WARNING

To trace the reasoning that drives us NUTS, we start as before. The NUT agrees with the MADman that it is essential to solve the credibility problem, but he favors a different solution. His plan is to find some sort of nuclear attack that would not only be a retaliation, but also would serve some vital purpose. Our threat would be credible because we would have, and we would be seen to have, a compelling reason to fulfill it. The retaliation we could have compelling reason to deliver is counterforce warfare. It is worthwhile to

destroy the enemy's forces, especially his strategic nuclear weapons. This reduces the risk to ourselves and our allies if war continues.

Thus we solve the credibility problem, and thereby we make it possible to succeed in nuclear deterrence – so says the NUT. But note a consequence of his argument: it has to be ambitious counterforce. If we want a highly credible warning that we would resort to counterforce warfare, there has to be little doubt that we expect its gains to be worth its risks.

But the drawbacks of an excellent counterforce capacity are these. First, and worst, an excellent counterforce capacity demands preemption. If our excellent counterforce capacity has been attacked, it may still be some sort of counterforce capacity, but it will no longer be excellent. The highly credible warning is, alas, not a warning of retaliation, but of preemption. Further, it gives the enemy his own incentive to preempt. His forces are under the gun: use them or lose them. Whatever use he may have in mind had better be done before it is too late.

This pressure to preempt is probably the gravest risk that the NUT embraces in his quest for credibility. But it is not the only one. Besides short-term instability in times of crisis, there is a second, long-term instability. The MADman could boast that his goals for deterrence are finite. Not so for the NUT. If we need enough capacity for counterforce warfare that we can credibly warn of our strong incentive to undertake it, then what we need is an increasing function of what the enemy has. In fact, we need superiority. For reasons the MADman has already given, a risk of arms racing is indeed a grave risk, both moral and prudential.

The third grave risk, this one primarily a moral risk, concerns the collateral damage from ambitious counterforce warfare. Given the proximity of missile fields to Moscow, for example, it makes little difference whether we target the population of Moscow per se, and so the NUT runs a grave moral risk of committing vast massacres, just as the MADman does. Not an equally grave risk: the MADman's attack is useless, whereas the NUT's is meant to destroy weapons that menace us. Further, the NUT's attack kills many fewer people. Too many people live downwind from the enemy's missiles, but not as many as live in the enemy's cities. Yet though the numbers that

measure the NUT's moral risk are much better than those that measure the MADman's, even the better numbers are far from good.

The MADman proposes to run grave moral and prudential risks so that a none-too-credible threat can be made very dreadful. The NUT proposes instead to run grave risks so that a somewhat less dreadful threat can be made very credible. His risks are different – most importantly, lesser massacres but more chance of inadvertent war – but no less grave overall.

EXISTENTIAL DETERRENCE

But what else can we do? How could the enemy be very powerfully deterred by a none-too-credible threat of a none-too-dreadful outcome?

This is how. Compare two ways a burglar might be deterred from trying his luck at the house of a man who keeps a tiger. The burglar might think: "I could do *this*, and then the tiger would do *that*, and then I could do *so-and-so*, and then the tiger would do *such-and-such*, and then. . . ." If all such plans turn out too low in their expected payoff, then he will be deterred. But if he is a somewhat sensible burglar, his thoughts will take a different turn. "*You don't tangle with tigers*. Especially when you've never tried it before. Not even if someone (someone you don't trust) claims that these tigers have somehow been tamed. Not even if you carry what the salesman claimed was a sure-fire tiger stopper. You just never know what might happen."

The hypothesis of existential deterrence is that it is through thoughts like these that our nuclear arsenal deters our somewhat sensible enemy. Existentialism says that the credibility problem more or less solves itself. Given an enemy who, like ourselves, is risk averse, pessimistic, skeptical, and conservative, deterrence is easy. To deter such an enemy, it is our military capacities that matter, not our intentions or incentives or declarations. If we have the weapons, the worst case is that somehow – and never mind why – we use them in whatever way he likes least. Of course he is not at all sure that the worst case will come about. But he mistrusts arguments to the contrary, being skeptical; and he magnifies the probability of the worst

case, being pessimistic; and he weighs it in deliberation out of proportion to the probability he gives it, being averse to risk. In short: he will be deterred by the *existence* of weapons that are *capable* of inflicting great destruction. And we are the same way.

If existentialism is true, then the package deals of MAD and NUTS fall apart. We can borrow ideas from the MADman and the NUT and have the best of both. But we can leave behind the parts of their reasoning that require us to run grave risks in order to solve the credibility problem.

BUY LIKE A MADMAN, USE LIKE A NUT

The MADman's policy for procurement of weapons was as moderate and benign as can be. The forces he requires are comparatively small and cheap. He creates no temptation to preempt. His standards of adequacy are finite, in the sense that both sides at once can meet them. We could be well content – if it were not for his abominable policy about what to do in case of war. But if existentialism is true, we can buy like a MADman if we like, but that implies nothing about what we ought to do in case of war, or what we ought to intend beforehand. We needn't strive to give some credibility to our dreadful threat to destroy the enemy's cities. We needn't threaten it at all. We have weapons and war plans which give us the assured capacity to do it, and their very existence is deterrent enough.

So far, so good; but a big question remains. What if we buy the MADman's finite deterrent, but it lets us down? What if deterrence fails after all, and in a big way? In that case, I say, we ought to use like a NUT. We ought to engage in counterforce warfare with what remains of our forces, hoping thereby to limit further damage to us and to our allies. We should not retaliate by destroying cities; on the contrary, we should compromise the efficacy of our attacks so as to reduce collateral death and destruction. We should proceed as if we valued the lives of the enemy's civilians and soldiers – simply because we *should* value those lives – but less than we value the lives of those on whose behalf we are fighting.

If we use like a NUT, but with nothing more than what remains of a MADman's forces, then our aims in counterforce warfare cannot

be too ambitious. We cannot hope to reduce the enemy's remaining forces to the point where he no longer has the capacity to do dreadful damage to whatever remains of our population and our resources for recovery. But the numbers count; they are not infinite, and not incomparable. If tens of millions are already dead, doubtless that is quite enough to exhaust our stock of adjectives and saturate our capacity to feel horror. But that is no reason why it is not worthwhile to save the lives of tens of millions more.

Limitation of further damage is worthwhile. Counterforce warfare, even of a modest sort, is a way to limit further damage. Therefore using our remaining nuclear weapons for counterforce warfare is the right thing to do. It is, of course, a better thing to do than destroying the enemy's cities. That alternative is easy to beat. But also, I say, it is a better choice than doing nothing, and waiting to see what sort of follow-on attack we suffer from the enemy's remaining forces.

It may be objected that it seems senseless to build forces designed for one mission when all the while we intend to use them only for another. If we buy like a MADman, we buy a force that is just right for retaliating against cities; but if the time comes to use like a NUT, we will wish the forces had been made more suitable for their only truly intended use: modest, second-strike counterforce warfare with avoidance of collateral damage.

Now it is the NUT's turn to have his package deal broken up. His policy about what to do in case of war – counterforce warfare meant to limit damage – is comparatively moderate and benign, at any rate compared to the MADman's. We could be well content – if it were not for his dangerous policy for procurement of weapons. Because he wants damage limitation not only for its own sake but for the sake of credibility, he requires weapons capable of meeting ambitious goals. Then the very same strength that supports the credible warning makes dangerous incentives to preemption in the short term and arms racing in the long term. Our solution is to buy suitable weapons, but limit their numbers.

Even a MADman's finite deterrent gives some significant capacity for counterforce. But all agree that the MADman's forces create little temptation to preemption or arms racing. They are not yet above the danger line. Then let them set a benchmark: let us have forces suited

for counterforce warfare, but let us have only enough of them to match the counterforce capacity of the MADman's finite deterrent. In that case, they should be no more destabilizing.

For finite counterforce, whatever enhances second-strike capacity without enhancing first-strike capacity is all to the good. Excellent post-attack command and control, for example, would be extremely advantageous. But it would not increase first-strike counterforce capacity in the least – because peacetime command and control is already excellent. Likewise, any improvement which holds capacity fixed while reducing collateral death and destruction is all to the good. If we aim our warheads more accurately and reduce their explosive yield (a trend that is already well under way), we can hold capacity fixed while we reduce the fallout, both local and global. And improved accuracy can mean that we need fewer warheads altogether.

If we trade numbers for accuracy, this reduces our capacity to destroy cities. Of course we do not have reason to want to destroy cities, but we do want the enemy to be deterred by the thought that somehow we might anyway. If the capacity is what deters, dare we reduce the capacity? I suggest that we can reduce it a lot without making existential deterrence any less robust. Any second-strike force that could accomplish something worthwhile in counterforce warfare, even with lower yields than we use today, would a fortiori be capable of enormous destruction.

CONCLUSION

As theoreticians, we want an understanding of nuclear deterrence that is neither MAD nor NUTS. We don't want to be committed to wickedness, and we don't want to fuss over credibility. We don't want deterrence through damage limitation – we want damage limitation for its own sake, and deterrence can look after itself. We don't want to think that damage limitation is worthless unless it is wonderful. We don't want to put adjectives in place of numbers, shirking the responsibility to save tens of millions of lives just because the outcome is dreadful either way.