

Chapter 8

**Making exceptions without
abandoning the principle: or how a
Kantian might think about terrorism**

THOMAS E. HILL, JR.

THE PROBLEM FOR KANTIANISM AND THE
LARGER ISSUE

Terrorism poses a practical problem that is urgent and difficult. How, within the bounds of conscience, can we respond effectively to violent terrorist activities and threats, especially given the ideological fanaticism and nonnegotiable demands that typically accompany them? The problem is partly instrumental and partly moral. The instrumental task is to find, among the morally permissible means, the best way to minimize terrorist violence, taking into account our other goals and values. The moral task is to determine what means of response are morally permissible. I shall focus here on this second problem, or rather on one way of thinking about it.

Terrorists, of course, often claim that their ends are morally worthy and that their means are morally justified in the context. Some of these claims deserve a serious hearing, and even the more outrageous claims can pose challenges that moral philosophers should not ignore. For present purposes, however, I shall simply assume that terrorism is morally indefensible, at least in the cases to be considered; and I will not be discussing why this is so. My inquiry, instead, is about what *responses* are permissible when terrorists *immorally* threaten the lives of innocent hostages.

Even this somewhat more limited question is too large to undertake here. To give an adequate answer would require

us not only to resolve deep issues in moral theory but also to investigate relevant matters of fact, make careful distinctions among cases, review our moral judgments regarding similar problems, and so on. Thus my remarks will address only one theoretical aspect of the larger moral issue: namely, how can a Kantian ethics, suitably interpreted and qualified, handle the life-or-death choices posed by terrorist threats? For example, can it ever permit the use of deadly force against terrorists even at the severe risk of killing innocent hostages? Can one refuse to negotiate for the lives of a few hostages in hope that a tough policy will in the long run save more people?

Although I shall explain the Kantian point of view I have in mind, my aim here is not to defend it either as a moral perspective or as an interpretation of Kant. Rather than presupposing the correctness of the Kantian perspective, I see my project here as preliminary to any final assessment of the relative merits of Kantian versus alternative perspectives: That is, it is simply an attempt to work out how a Kantian approach might direct us to think about some of the terrible choices forced upon us by terrorists. More specifically, can a person (rather like me) who is sympathetic to some main features of Kant's ethics make (intuitively) reasonable judgments about terrorism without, in effect, abandoning the fundamental Kantian principles? And if one admits that terrorism calls for "exceptions" to the straightforward applications of Kant's basic principles for normal circumstances, can one acknowledge these exceptions coherently, and not in an ad hoc manner, from a moral perspective that preserves much of the spirit of Kant's ethics?

My project, I should stress, is not to use Kantian principles to draw a precise line between permissible and impermissible responses to terrorism but rather to see whether a basically Kantian ethics has any reasonable and coherent way to approach the troublesome cases. Many of these cases leave me morally perplexed. I am not confident about what would be right to do. But even this uncertainty is enough to raise the problem, for it seems at first that basic Kantian principles

leave no room for what generates the uncertainty. What pulls us even to consider, say, sacrificing a few to save many seems to be a kind of reasoning that makes sense only within a moral perspective radically opposed to Kant's. Even to acknowledge the reasoning in favor of the exceptions, it seems, is to abandon the Kantian perspective.

The issue here is an instance of a more general problem faced by many people who have never heard of Kant but who believe in taking a "principled stand" on moral questions. The problem is posed sharply by an old dialogue that could take many forms.¹ One person, A, asks another, B, whether for a million dollars B would do something against B's principles (e.g., go to bed with a stranger, tell a racist joke, or take hotel towels without paying for them). B responds, "Yes, I suppose I would." Then A asks, "How about five dollars?" B retorts indignantly, "What do you take me for!" "We have already established that", says A. "Now we are dickering over the price."

To those who take pride in standing on principle, the story poses a dilemma: Either you hold rigidly to your principles for all circumstances or else you are prepared to suspend them, or open gaps in them, for special cases. In extreme abnormal circumstances the first course may begin to appear foolish, contrary to both common sense and reflective judgment; but taking the second option raises the suspicion that one's principle is just a rule of thumb and that one is open to persuasion by considerations that have nothing to do with the rationale for the alleged principle.

The sort of dialogue just illustrated was not meant to pose a serious moral dilemma, but more dramatic cases are not hard to imagine. Suppose, as in John Fowles's *Magus*, one were forced to choose between beating two Nazi resisters to death and having the vicious Nazi commandant shoot down a dozen or more resisters. Suppose one had to torture a terrorist in order to learn the location of a bomb set to blow up New York City? If we admit exceptions to familiar moral principles in these cases, then critics can start to push us down the slippery slope. How about if the bomb would blow

up only the Bronx? Or a half of that? Or a dozen people? Or two? If we take a hard stand on principle, refusing to make the exception, then the critics' move will be to "up the ante," asking, for example, whether we would kill the two resisters or torture the terrorist in order to save the whole world? The critics' objective here is to open a hole in the absolutist dike, leaving only consequentialist resources to hold back the flow of exceptions. They may even admit that it is almost always right to do as the principle prescribes, for the main issue is not how frequent the exceptions are. Rather, the critics want to raise a suspicion that there is no good reason to maintain the principle *as a principle*, as opposed to a flexible rule of thumb.²

One who insisted on standing by ordinary principles in all circumstances, despite the critics' challenge, could of course take a leaf from the consequentialist's book, denying the relevance of merely fictional cases and arguing that we do not actually face such stark alternatives. But though options in the real world are rarely as clear-cut and certain as those in philosophers' examples, to insist repeatedly that each tough case is merely fictional seems a lame defense, one that appeals more to faith than to experience. Besides, since defenders of principle have so often used fictional counterexamples against utilitarianism, they should be more than a little embarrassed to resort to the wholesale rejection of hypothetical examples when trying to uphold their own position.

The history of ethics is full of subtle attempts to meet our imaginary critics' challenge to those who hold familiar moral principles as principles. I shall not review these attempts here, but one common strategy deserves special mention. This is the tendency to append a "catastrophe clause" to familiar principles whenever the consequences of adhering to the principles are so repugnant that it seems morally perverse to refuse the exception – for example, "Do not even threaten to kill innocent people, unless doing so helps to avert a nuclear war" or "Let justice be done, unless (thereby) the heavens should fall." This strategy of adding catastrophe

clauses formally maintains the principle as nearly inflexible and yet allows us to side with common sense on extreme counterexamples. The problem with the strategy, however, is that it leaves our critics' main challenge unanswered: If the balance of consequences determines what to do in the extreme case, why not in the case slightly less extreme, and so on? If one may sacrifice a life to save a million and one lives, then why not to save a million, or a million minus one, or a million minus two, or . . . three, or four . . . ? Isn't it basically a matter of calculation, after all? This remains a serious challenge unless one can give a coherent common ground both for maintaining the familiar principle as nearly inflexible and for making exceptions for extreme cases.

Although my discussion will be somewhat abstract and theoretical, the sort of problem raised here is not merely a philosopher's puzzle but a crisis of conscience for many people. Try to imagine, if you can, one of my typical undergraduate students whose talents and ambitions lead her into the inner circles of Washington facing decisions and a world view more akin to those of Oliver North than to those of the local minister. Raised in a small peaceful community with shared values and limited global outlook, she has come to take for granted certain absolute prohibitions as the framework of any tolerable social life and of any decent person's conscience. Unreflective religious beliefs tend to reinforce her principles, but one suspects that it is not so much that her principles are derived from religion as that her theological beliefs are reinforced by the way they wrap an aura of authority around those principles. In the local world the hard cases did not force themselves upon her thinking. Removed to the tougher world of international conflict, power politics, terrorism, and constant compromise, however, she must live with the widespread opinion that her initial principles are foolishly rigid and face choices where sticking to those principles seems unbearably costly. Making an exception here and there, perhaps with a bad conscience, may not at first force her to rethink her position; but after a while she may well wonder how she can continue to make what seem at

each time justifiable exceptions without having, in effect, abandoned her principles. Ships with too many holes sink, and principles with too many gaps no longer function as principles.

KANTIAN PRINCIPLES AND THE PRESUMPTION AGAINST KILLING

For the would-be Kantian the challenge takes two forms: First, can the Categorical Imperative, in some form, be sensibly maintained as an inflexible principle in the face of extreme cases? Second, if one admits intuitively reasonable exceptions to the more specific principles about killing (lying, promise keeping, etc.) commonly associated with Kantian morality, in what sense can these still be maintained as principles without abandoning the basic Kantian point of view? For example, can the exceptions themselves be justified by appeal to the Categorical Imperative, or are they, after all, partial concessions to a radically different moral point of view? My main focus will be on the first question, but what I say will also be relevant to the second.

Unlike many who are sympathetic to Kant's moral philosophy, I have little confidence that Kant's famous "universal law" formulations of the Categorical Imperative can function adequately as guides to moral decision making. More promising, I think, is Kant's principle that humanity in each person must always be treated as an end in itself, never simply as a means. This principle, Kant says, is unconditionally binding on all human beings, whatever their circumstances and regardless of what (contingent) ends must be sacrificed to satisfy it. It is, supposedly, not merely one "perfect" duty among others but an articulation of the comprehensive ground of all duties, "perfect" and "imperfect." As a foundation of a rational system of moral principles, it is not supposed to generate or allow any genuine conflict of duties or (what are often called) "moral dilemmas." If unresolvable conflicts of duty are derivable from it, then it cannot have the status Kant claimed for it.

Now despite the warm reception Kant's principle receives from nonconsequentialists everywhere, its interpretation remains controversial. For our purposes, however, I shall simply take as given the following reconstruction.³

First, the "humanity in persons" that we are urged to respect is the "rational nature" of human beings, or human beings insofar as they are considered as rational agents with autonomy. The rationality and autonomy in question here are capacities and dispositions that virtually all sane adult human beings are presumed to have, not the full manifestation of these in actual conduct. Thus people who act foolishly and immorally are, in the relevant sense, rational agents with autonomy, and their "humanity" too must be treated as an "end in itself."

The status of infants and the mentally incompetent remains problematic under the principle, and the principle does not address the value of nonhuman animals. But we can set aside these special problems if we limit our examples to mentally competent, adult human beings and do not presume that the humanity principle alone is adequate to settle all moral issues.

Second, the injunction not to treat humanity merely as a means is incomplete and inapplicable without the contrasting phrase "always as an end in itself"; for without knowing what it is to treat humanity as an end in itself we cannot judge what would be, in the intended sense, treating it *merely* as a means. In contrast, if we succeed in treating humanity always as an end in itself, we automatically satisfy the prohibition against treating it merely as a means. Further, as Kant makes clear in his later writings, treating a person's humanity as an end in itself requires more than refraining from acts that would exploit the person as a mere means. Indifference to a person is also forbidden, and positive assistance may be required.⁴

Third, the term "end" here, as Kant acknowledges, is a potentially misleading technical term, meaning roughly something giving a reason or "ground for determining the will." An "end in itself," or "objective end," is a ground of choice that would determine the will of a fully rational agent and

should determine ours. In the ordinary sense, of course, neither persons nor the rational nature of persons is an "end."

Fourth, in saying that rational agents with autonomy are ends in themselves, I take it, Kant is saying that they have "dignity," which he defines as an "unconditional and incomparable worth" that, unlike "price," "admits of no equivalent." As an unconditional value, dignity does not depend on the contingent fact that something is useful, desired, or even liked. By contrast, things with mere "price," or conditional value, have a value that is dependent upon utility ("market value") or at least upon individual sentiments ("fancy value"). Anything with mere price has a value that "admits of equivalents" and so is subject to calculated trade-offs; however great its value, there can in principle be something else that could compensate for its loss and justify its sacrifice. Material goods, reputation, and pleasures as such have mere price, and so even great amounts of these things may at times be reasonably sacrificed for other things with the same sort of value. By contrast, dignity is "above all price" and so one can never act contrary to the dignity of someone for the sake of things with mere price, no matter how great the price.⁵

Fifth, Kant's fundamental principle is directly concerned, not with "external actions," but with the attitudes, or value priorities, that lie behind the ways we treat ourselves and others. Acts, for example suicide and murder, and institutions, for example slavery, are condemned not merely as kinds of intentional behavior with undesirable consequences but as manifestations of value priorities that are intolerable. The example of suicide is instructive here. In the *Groundwork*, for example, the suicide Kant condemns is terminating one's life because a hedonistic calculus predicts that one's future will contain more pain than pleasure.⁶ Kant later says that to avoid suicide is a "perfect" or exceptionless duty, but he admits that it is an open question for casuistry whether we should *count as suicide* deliberate self-sacrifice for one's country or killing oneself to avoid an unjust death sentence or terminal madness from the bite of rabid dog.⁷ The operative principle here is obviously not that intentional killing of a

human being is always wrong but rather that (with a few conceivable exceptions) the reasons people have for killing themselves have only a conditional value ("price") that is not to be compared with human dignity.

This understanding of Kant's dignity principle fits well with Kant's other uses of the idea of humanity as an end in itself in the second part of *The Metaphysics of Morals*, which is concerned with ethics beyond questions of law, justice, and rights. For example, the principle is used to condemn drunkenness and gluttony (which interfere with rational functioning), mockery (which treats people as worthless), and servility (which expresses the attitude that one's human worth is measured by one's utility to others).⁸ The duty to respect others is grounded in the value of their humanity, not in their achievements or their moral conduct; and significantly, the duty to promote the happiness of others is not grounded explicitly in the dignity principle.⁹ In summary, these principles say that one must seek to preserve, develop, exercise, and "honor" rational agency in oneself and to respect it in other human beings, no matter how immorally and irrationally they may behave. To preserve human life per se is not among the principles.

Now though these applications illustrate the dignity principle, they are not concerned with the sort of life-or-death choices that terrorism can force us to confront. Regarding these cases, what Kant says about law and justice, in the first part of *The Metaphysics of Morals*, is more relevant. But before we turn to that, we need to face a further question about the interpretation of the dignity principle.

In brief, the problem is this: Granted that dignity cannot be exchanged for price, can dignity be exchanged for more dignity? To put the question less cryptically: Assuming, as already decided, that dignity is "above all price" and so cannot be sacrificed for things of merely conditional value, does the principle imply, further, that dignity itself is a nonquantitative value that does not admit comparisons like "the dignity of this (person) is greater than the dignity of that (person)" or "two things of dignity are worth twice as much

as one"? Does dignity merely "admit of no equivalent" among the things of conditional value or are there no equivalents even among things of dignity? Since dignity is the value attributed to human beings, these questions obviously have a bearing on the question when, if ever, the dignity principle permits the sacrifice of one human being for the sake of one or more other human beings.

Now when we read that Kant approved of capital punishment, and sometimes killing in war, it is tempting to jump to the conclusion that Kant himself understood the dignity principle in the more permissive way, allowing us to justify some killing of human beings on the ground that the dignity of many persons outweighs the dignity of a few. However, for textual reasons I shall pass over here, I do not think that this interpretation is correct. More important for our purposes, the permissive interpretation would undermine what seems to be the strikingly special character of Kantian ethics, namely, its refusal to reduce moral decisions to the weighing and balancing of commensurate values.¹⁰ On the permissive interpretation, the theory would become just another version of an all-too-familiar type: namely, a theory that first assigns a quantity of intrinsic value to various possible outcomes and then treats the right decision as a function of these value assignments. Kant's theory, to be sure, would have the unique feature of having two scales of value (dignity and price), with any amount of value on one scale always "trumping" any amount on the other; and it would place the higher value on living as a rational agent as opposed to hedonistic values. But nonetheless, like sophisticated forms of utilitarianism that place a higher priority on human life than on animal pleasures, such a theory would hold "the good prior to the right" and make moral choices fundamentally matters of calculation.

Thus, to further discussion, let us work with the alternative, nonpermissive interpretation. On this account, dignity is still "above all price" in the previous sense: Dignity must never be violated in exchange for things of mere price. But now dignity is not to be construed as a quantitative notion.

Dignity is "without equivalent" even among other things with dignity in the sense that one cannot justify a violation of dignity by claiming it as a necessary sacrifice to promote "more" dignity elsewhere. Thus the destruction of something with dignity, if ever justifiable, cannot be justified simply by weighing quantities of intrinsic value, even value of the highest order. The assignment of dignity to each rational agent, then, functions not to introduce a new kind of value calculation, but rather to block our tendency to treat rational agents as interchangeable commodities. Moreover, "the right" remains "prior to the good" in that the attribution of dignity to rational agents stems from a rational/moral command to adopt and act consistently with a certain attitude toward rational agents, not from a metaphysical claim about intrinsic values or from an empirical claim about preferences fixed by human nature.

Though the analogy is not perfect, we may say that the Kantian attitude toward preserving the lives of rational agents is rather like the attitude that some pious persons take toward an object they regard as "sacred" or "holy." They revere and treasure the object and seek diligently to preserve it. Even the thought that they might have to destroy the object is abhorrent, and so they make every effort to avoid situations where that would be necessary. They view any such object as "irreplaceable," not merely in the sense that they would not trade it for an exact copy but in the sense that they would not view even its replacement by many other things revered in the same way as commensurate values compensating for the loss. Nonetheless, if certain deplorable circumstances arose, they might grant that they must destroy the object, for example, to keep it from being defiled by enemies. The attitude that such objects are "without equivalent" does not translate simply into an action principle that they must never be destroyed but is more complex. It is also concerned with how one should view their loss, how one must work to prevent situations where losses will inevitably result from one's choices, and how one should think about what to do when such tragic situations occur.

The attitude in question is perhaps more intuitively recognizable when one reflects how a loving parent would view the horrible situation of being able to save one of her children from drowning or to save two others, but not all. Surely she is justified in saving the two, but it is hard to conceive that she would accept that the rationale was that two are worth twice as much as one. The problem, of course, is to say what other rationale there can be.

It is the more radical, nonpermissive reading of the dignity principle, I believe, that accounts for its wide appeal; but this is also what raises most acutely the consequentialist's challenge. The principle does not absolutely prohibit killing human beings, nor does it necessarily forbid us, when faced with no other option, to make the hard choice that results in the death of a few and the preservation of many. But if such choices are morally defensible, they cannot be rationalized simply by the thought that "the lives of many people are worth more than the lives of few people" or even "many innocent lives are worth more than fewer." Even more obviously, the dignity principle does not permit the choices to be justified by the thought that by preserving the lives of more happy people we will bring about the most "intrinsically valuable" experiences.

The problem, then, is how, despite this severe restriction on consequentialist thinking, can the Kantian ever justify making the hard life-or-death choices that seem intuitively compelling, particularly in cases where at some level of deliberation the numbers do seem to matter? The problem is far easier to state than to answer, but some clues toward an answer, I think, can be found in Kant's views on state coercion and punishment, to which I now turn.

DIGNITY, PUNISHMENT, AND DEADLY FORCE AGAINST TERRORISM

Kant held that justice not only permits but even requires capital punishment for murder. I disagree, but I think that what Kant says in support of his view at least suggests how

the dignity principle can be reconciled with the use of deadly force against terrorists in some cases. It also opens up a way of thinking about cases where a response to terrorists would endanger the lives of innocent persons. Once again I shall reconstruct Kant's views liberally, without argument, but the purpose here is to focus attention on the possible reconciliation, not on questions of textual interpretation.¹¹

The background assumption is that we are dealing with agents who are rational though not necessarily moral or ideally rational. They know what they are doing, are capable of self-control, and know well enough the difference between right and wrong. Moreover, they are disposed, even when they act immorally, to acknowledge moral requirements as rational and not merely imposed by others. When conscience and moral argument fail to dissuade them from their immoral projects, they are at least rational enough, as a rule, to be deterred by clear and credible threats to their own welfare. These assumptions, of course, are not always satisfied in the real world; and to the extent that they are not, the Kantian rationale may be vitiated.

Next, let us grant that, for various reasons, human beings need to join together into communities with common laws that assign to each rights and responsibilities. In a state of nature (or anarchy) individual reason and conscience would not in fact suffice to create and preserve the conditions necessary for people in close proximity to have lives appropriate to rational agents. So states are needed to define and secure to each person a reasonable opportunity for life and liberty as a rational autonomous agent. Thus there are moral as well as prudential reasons to form and support civil authorities with coercive powers.

This being so, the dignity principle must be applied not merely to relations between individuals isolated from others but, first and crucially, to the construction of a system of laws that can provide the framework for moral relations among individuals.¹² In working out the implications of the dignity principle for laws, one should take the point of view of legislators in the "kingdom of ends," that is, fully rational

and autonomous persons, each with "private ends" but "abstracting from personal differences," and making only universal laws.¹³ From this point of view, Kant and Rawls agree, the first principle of justice adopted would be that the system of laws should try to ensure to each person, viewed in advance of particular contingencies, an equal and full opportunity to live out his or her life as rational agent within the constraints of those laws. This is essentially what Kant calls "the universal principle of justice," and, of course, it is more or less Rawls's first principle of justice.¹⁴

Unfortunately, human nature being what it is, merely articulating laws that coordinate activities and define fair shares of liberty to each will not suffice to ensure compliance. Many would overstep the bounds and encroach on others' liberty. Thus coercion is needed as a "hindrance to hindrance of freedom."¹⁵ The coercive power of the state must provide incentives so that even without conscience everyone will have clear and sufficient reason not to violate the liberty of others (as defined by just laws). The aim of the coercive legal system is not to maximize welfare or to give the morally vicious their just deserts but rather to create the conditions in which each has, so far as possible, a fair chance to live out a life as a rational autonomous agent.

State officials cannot, of course, be present at each moment a person begins to invade someone else's freedom contrary to just laws, and so a system of punishment is needed. Though applicable only after a crime has been committed, it can serve to "hinder hindrances to freedom" by credible threats that provide rational incentives, apart from conscience, for each to stay within the bounds of the freedom he or she has been fairly allotted. The threats must be genuine, enforceable, and public in order to be credible, and they must be carried out as legally prescribed for the sake of both fairness and efficacy.

The system of punishment, however, must not only be a rational deterrent; it must also honor the dignity of each rational agent. This means that criminals must be treated with respect as human beings, not humiliated or manipulated

like animals. It also means that the criminal, like every other citizen, must have had a fair chance to avoid the penalty. Thus *ex post facto* laws, strict liability, and unpredictable penalties must be avoided; the criminal law must be public, easily understood, uniformly enforced, and concerned with provable "external actions" rather than inner moral qualities. Ideally, the penalties would be just severe enough so that, given effective enforcement, reasonable self-interested persons could see that each crime is unlikely to be profitable and yet standards of proportionality of penalty to offense are maintained. The system must be defensible even to the criminal insofar as he or she is willing to look at the matter from the perspective of one rational agent, with dignity, among many.

Kant thought that the system that best satisfies these criteria relied on the *jus talionis*, or "an eye for an eye," as the rule for determining the degree and kind of punishment that should be attached to each offense.¹⁶ Thus execution should be the punishment for murder. As I have argued elsewhere, this is not because the "inner viciousness" of the murderer deserves the death penalty, still less because the murderer has forfeited all consideration under justice and as a human being. Nor, I think, is Kant's main argument the implausible claim that in choosing the crime with a known penalty the criminal has actually willed his or her own punishment. The rationale implicit in the central features of Kant's theory of law is that capital punishment is a necessary part of a fair system needed to secure to each citizen, so far as possible, the opportunity to live out a life as a rational and autonomous agent.

We may doubt various assumptions, for example, that only the death penalty can provide adequate (nonmoral) disincentives to murder and that fairness requires inflexible application of the penalty. But let us grant these points for now. How does the reconstructed rationale for judicial killing of human beings square with the dignity principle? That principle, applied to legal systems, clearly implies that criminals should be treated with respect, that the death penalty should

be imposed reluctantly, and that the systems should contain other features designed to prevent the crimes that call for capital punishment. The dignity principle also implies that capital punishment cannot be justified by the argument that it saves money, reduces fear, or promotes the happiness of most people. Nor can the rationale be simply "we will save more people than we kill" or "the people we save are worth more than those we kill."

But the justification I sketched was not of these kinds. It used the dignity principle to set standards for a general system of laws, concluding that a legal system expresses its respect for the incomparable worth of each rational agent by seeking to secure for each, in advance of particular contingencies, a full fair opportunity to live as a rational agent. A general system of laws acknowledges the worth of citizens, not as defined at a given time by character traits, social roles, achievements, and records of conduct, but as conceived more generally as rational agents extending over time, capable of choosing any number of roles and courses of action. In carrying out just punishment, one can argue that the system did, so far as possible, secure for even the criminal, in advance of the criminal's particular choices, a full life with as much liberty as possible in an impartial system of laws. Though the criminal does not wish for the courts to apply the punishment in his or her case, the system respects as far as possible what he or she wills, or would will, as a system of laws when looking at the matter just as a rational agent apart from a particular history and circumstance. If, as Kant thought, capital punishment for murder is a necessary part of that system, criminals should have no legitimate complaint that the law failed to respect their human dignity.¹⁷

This line of thinking, if cogent, can readily be extended to justify laws authorizing the police in some extreme cases to kill a terrorist who is immediately threatening the lives of law-abiding citizens. By my initial assumption, our terrorist, like the murderer, is a rationally competent agent who flagrantly and immorally crashes across the boundaries of freedom that a just system of law tries to ensure for each,

including him- or herself.¹⁸ By authorizing police to use necessary force, even deadly force, against terrorists actively engaged in life-threatening activities, the system serves to "hinder hindrances to freedom" without violating anyone's dignity. In fact, the case here for deadly force against terrorists in action seems stronger than the case for capital punishment for murder because in the latter case prevention efforts have already failed and it remains controversial whether capital punishment is necessary as a deterrent.

DIGNITY AND ENDANGERING INNOCENT HOSTAGES

So far I have considered only how the dignity principle might be reconciled with policies that involve taking the lives of the terrorists themselves, given some strong assumptions about what the terrorists are like. This suggests that a Kantian may admit legitimate exceptions to the rule "Do not kill human beings," even though there remains a strong presumption in favor of that rule in most circumstances. More controversial questions arise, however, when we consider responses to terrorism that endanger the lives of innocent people. In these cases we could not argue to the persons whose lives we endanger that they have knowingly overstepped the bounds of liberty allotted to them under a fair system of law. Moreover, in at least some of these cases common sense tells us that the numbers matter, that (for example) risking the life of one hostage to save hundreds is not unreasonable.

The issues here are complex. Many variables may be relevant. For example, do we have an option to negotiate without endangering anyone? If not, will the deaths result from our violent attack on the terrorists or from the terrorists themselves when we refuse to negotiate or attack? Are the persons we put in jeopardy already in danger from the terrorists? Is there a chance of rescuing the very persons we put in danger? How do we estimate the odds of our killing the hostages, of the terrorists' killing the hostages, of our rescuing the hos-

tages, and of the terrorists' killing other people? And how certain are we of these estimates? How many survivors and how many dead do we expect in each scenario? How pure are our intentions in imposing the risks? Is there publicly known and accepted policy concerning these situations?

We cannot review all variations, but let us begin with one of the easier cases. Suppose that negotiation to free the hostages is not an option, and so we must attack or stand by and accept the consequences. Suppose, further, that it is nearly certain that all the hostages will be killed if we do not make a rescue attempt. If we do make the attempt, there is a rather good chance that we can save all the hostages, but there is also a nonnegligible chance that we may kill some hostages in the effort. For simplicity, suppose that the danger to the attackers and other innocent people is minimal.

Most people, I suppose, would say that we should make the rescue attempt and that if, despite our precautions, we kill innocent hostages in the effort we were still justified in doing what resulted in their deaths. After all, they might say, it is a question of some being killed or all being killed. So, it seems, the intuitive judgment is that the numbers matter here. Since the lives of nonthreatening, law-abiding persons are at stake, the Kantian cannot accommodate the intuitive judgment by the same argument we considered for capital punishment and for using deadly force against the terrorists themselves. The Kantian must refuse to permit the attack, contrary to common sense, or else find a new argument.

The case is made easier by the fact that we stipulated that hostages who may be killed in the attack would be killed by the terrorists if we do not attack. We do not need to endanger one person in order to save others. We need only to justify risking our killing that person in an effort to prevent that person's nearly certain death at the hands of others.

Could the Kantian ever justify this? Again, the dignity principle implies a reluctance to kill or even risk killing human beings, and it implies that we should make every effort not to be forced into situations where these actions might be

required. But as the dignity principle does not absolutely prohibit killing human beings, it cannot absolutely forbid taking a serious risk that we may kill human beings. The question is, can we risk killing innocent persons in the sort of situation at hand?

Taking a clue from the argument for state coercion and capital punishment, the Kantian should consider the question not in the isolated instance but as it falls under general laws and policies applicable to everyone over time. Here the issue is: What public policies would it be morally legitimate for a community to adopt in authorizing official responses to terrorism? As before, the appropriate Kantian point of view for deliberating about this is presumably that of a rational legislator in the "kingdom of ends."

The legislators' commitment to the dignity principle means that they have a strong concern to preserve each person's opportunity to function as rational autonomous agent. So they have an initial interest in saving lives. But the dignity principle also imposes constraints on how they can justifiably do this. For example, the legislators cannot simply assign a quantity of value to each life and then try to find the value-maximizing policy. Moreover, they must find policies that are justifiable even to those who will suffer under them (insofar as they adopt the same legislative point of view). Further, since human dignity means more than merely preserving life, the legislators must consider other ways that policies can affirm that central value. They would, for instance, tend to favor policies that strongly express and encourage mutual respect, and oppose the opposite, independently of the policies' effects on survival rates.

Considerations of dignity alone are not always sufficient to resolve an issue. But nonetheless there may be a reasonable way of thinking about the problem from the Kantian legislative point of view. Each legislator not only values human dignity but also has a set of "private ends" that he or she can pursue only while living. (I set aside here beliefs about immortality.) Valuing each other person as an end, the legislator also has reason (Kant implies) to give some

weight to the ends of others. From the legislative point of view lawmakers must "abstract from the content of private ends" and so cannot concern themselves with the details of who wants what. But that each has a general concern for his or her own survival seems to follow from the idea that members have private ends (and know so). To be sure, legislators are committed to seeking policies promoting their own survival only on the condition that each policy is equally acceptable to others. But others have a similar concern for their own survival. Thus, one might argue, all would favor any policy that promised to prolong the survival of representative persons as rational autonomous persons, other things being equal and provided the policy is otherwise consistent with the dignity principle. Each legislator would favor the policy when looking at it from his or her own perspective (abstractly conceived) and also when reviewing it from the perspective of each other person.

Now if suitably hedged with restrictions, some policy that permitted authorized agents to endanger the life of a hostage in an attempt to save that hostage and others from almost certain death would be seen, in advance of particular contingencies, as enhancing each representative law-abiding citizen's chances for surviving terrorist situations. If so, the policy would be reasonable and, it seems, not inherently contrary to the dignity principle. If, in addition, the policy was decided upon in a legitimate political process, was known to all, and was carried out by authorized individuals, then (absent further objections) the dignity principle would seem to permit its execution, even at the serious risk of innocent lives.

This argument is too sketchy to draw any precise line between permissible and impermissible cases, but it suggests at least that the Kantian is not forced to choose between treating dignity as a measurable quantity and condemning all avoidable killing of innocent persons. The reasoning allows us to consider "chances of survival" and so (indirectly) the numbers of lives at stake, but this consideration enters the deliberation not as a basic principle ("many lives have

greater value than few") but as pertinent information at a later stage. "Try to save the most lives" may become a highly qualified derivative principle for specified circumstances, but this is compatible with the dignity principle. And of course, to deny that the chances of survival under different policies should ever in any way be a relevant moral consideration would be sheer madness.¹⁹

A more difficult question arises when the persons endangered by our efforts to free the hostages are not the hostages already at risk but others. Even if the attack will save many lives, it cannot be defended as giving a better chance of survival to the very people endangered by the terrorists. To attack, it seems, would be to treat the fewer persons as exchangeable for a greater number of persons, contrary to the dignity principle.

Although I think there are strong reasons not to endorse a general permission to attack in such cases, the Kantian line of thinking we have been considering need not lead to an absolute prohibition. It might allow the attack in certain carefully circumscribed extreme cases but, of course, not on the ground that "more lives are worth more than fewer." The reasoning in favor of the attack, as before, would have to be that the policy of attack in exceptional cases of this kind could be defended from the general legislative point of view even to those who turned out to be the victims under the policy. The argument from that point of view would be that the policy, all things considered, enhanced each representative (law-abiding) citizen's chances of survival, without violating any of the essential constraints of the Kantian legislative perspective. To show this we might need to establish that the chances of each citizen's falling into the unfortunate hostage role, independently of his or her choice to take special risks, is more or less the same. We would also need to take into account the likelihood that the policy would undermine respect for life, would lead to foolish risk taking, and would be abused by callous, trigger-happy officials. A legal system or moral code that said, without careful qualification, "Authorities may risk killing one person whenever they think

that they can thereby save more than one" would invite abuse and probably cost more lives in the long run. Moreover, prolonging human life is not the only value implicit in the dignity principle. Any policy must also be evaluated according to the way it expresses and encourages mutual respect, honors prior commitments, upholds just institutions, and so on.²⁰

Nonetheless, a more circumscribed policy of this sort could conceivably give each representative citizen a better prospect for living out a full life as a rational agent than would an absolute prohibition on endangering innocent lives in such cases. Though the stakes may be higher, the argument for the policy would in principle be like the case for current laws that permit police to speed when actively pursuing armed robbers even though this imposes risks on law-abiding motorists. If all other constraints were satisfied, such policies might be approved by all from the appropriate point of view.

In this way of thinking, our calculation of each representative person's chances of survival under different general policies would be information relevant in the overall justification, but the fundamental principle regarding official responses to terrorism would remain "Public policies must conform to principles acceptable to all from the abstract legislative point of view that, among other things, regards each person as of incalculable worth." One way of respecting the dignity of all is to authorize police to impose grave risks only as permitted under public policies that they themselves, as rational agents, would approve.

INTENTIONAL SACRIFICE OF INNOCENT PERSONS

My aim in reviewing this series of cases has been to suggest a (more or less) Kantian way of thinking that might reconcile the dignity principle with the common opinion that in some circumstances it is morally permissible to kill terrorists and to endanger innocent hostages. The scenarios we reviewed posed questions only about taking risks that we will kill innocent persons, unintentionally, as we try to save others.

Reasonable people may differ about the degree of risk that is warranted, but almost everyone agrees that some such risks are worth taking. Unfortunately, we can imagine situations where terrorism confronts us with choices more horrible than those we have considered, and about these cases philosophers will disagree more sharply. For example, suppose that the only way to prevent a terrorist from blowing up a building with many hostages in it is to shoot the terrorist through another innocent hostage who is held as a shield. Or, still worse, suppose that we could save the many hostages only by deliberately killing an innocent victim of his choice.

Here traditional morality draws the line, refusing to authorize the intentional sacrifice of an innocent person even to save others.²¹ Kant too would be horrified at the thought, and no doubt many consequentialists would feel the same way. But the question here is whether the theories in question could permit the killing, and how they could justify their conclusions.

These more extreme cases pose a challenge to the Kantian quite different from the one that has been my main concern. The worry now is not that the Kantian cannot permit *enough* of what reasonable persons would allow; it is that the Kantian *as construed here* will permit *too much* to be morally tolerable, authentically Kantian, or significantly different from the consequentialist.²²

These are important concerns, deserving a fuller treatment than I can attempt here.²³ A thorough response would require a more detailed characterization of the Kantian legislative perspective than what has been sketched here. But even without that, it seems clear that our Kantian legislators have strong reasons for refusing to authorize the intentional killing of innocent persons in response to terrorism.

Most obviously, all the considerations making the Kantian legislators reluctant to *endanger* innocent hostages will be even stronger reasons not to authorize *intentional* killing of innocents. To legitimize any such killing as public policy would be a dangerous invitation to abuse by corrupt, self-

deceiving, and bungling officials. The policy would likely lower everyone's sense of security and trust. Those in authority could not honestly say even to a loved one, "I would never kill you or turn you over to killers." Worse yet, official readiness to kill innocent persons in response to terrorists' threats not only would encourage such blackmail but also would enhance the terrorists' power. By making convincing threats to do a greater evil, terrorists could easily manipulate the authorities into doing a lesser one. The lawless could thus enlist the aid of law enforcement to get their dirty work done for them.

Few would deny that the general policy of killing innocents to placate terrorists would have these bad consequences, and more; but some may try to press an objection often posed to consequentialists. That is, they may argue that, in some extraordinary situation, a one-time, secret, unauthorized sacrifice of an innocent person might save lives without incurring all the problems of openly admitting a policy of negotiating with terrorists. Here, the objection continues, the killing would still be wrong but not for the reasons given above.

Now this counterexample may pose a problem for consequentialism of some kinds but is irrelevant to the Kantian legislative thinking proposed here. This idea was introduced not as a comprehensive moral decision procedure but as a thought experiment to guide and constrain attempts to justify public policies with regard to issues, such as murder and terrorism, where not having enforceable public policy would be disastrous. A crucial background assumption was that the procedure justifies policies only *as openly acknowledged and scrupulously administered*.²⁴ If any "exceptions" are to be justified by the legislative procedure, then, they must be justified as overt aspects of a general public policy, not as secret, one-time deviations.

A Kantian ethics must also take into account the following. The dignity principle requires that we *always* regard each person as having incomparable worth, and so not as a mere means even to worthy ends. This is an attitude to be main-

tained in daily living as well as a stipulated constraint on thinking from the abstract legislative standpoint. Thus, to be justified in a deliberate killing, a person would need to be able to face the victim and say, *sincerely and truthfully*, "I choose to kill you (when I have an option not to); but still I regard you as more than a mere means, in fact, as a person with a worth that is incalculable."²⁵ Now it is not easy to imagine that a human being could *maintain* this attitude while deliberately killing another, even though apparently Kant thought that public executioners could and many think that doctors administering euthanasia can.²⁶ It seems even more unlikely that human beings in general could authorize, approve, and carry out public policies of deliberate sacrifice of innocent persons while continuing honestly to affirm an untarnished respect for human dignity. As Kantian legislators they could not approve of any policy that they could not carry out consistently with the dignity principle. Thus it seems implausible that they could accept any policy involving deliberate sacrifice of innocent persons even if they predicted that it would save more lives.

RESPONSIBILITY AND "CLEAN HANDS"

A final issue that calls for fuller discussion but can only be briefly mentioned here is how we should determine responsibility for the deaths in our terrorist situations. The worry is that our abstract deliberations about the numbers of lives saved and lost have not taken into account *who* will be responsible for the deaths. Mustn't we consider not only how many will die but also whether the deaths result directly from our acts (or omissions) or from another's immoral response to what we do (or refuse to do)?

Traditional morality does not impute to me, as the "consequence" of my action or inaction, every event to which my action or inaction was a causally contributing factor, even if that event would not have occurred but for my action or inaction. We are "responsible" for some of the effects of what we do but not for others. If, for example, prior to President

Bush's inauguration someone had threatened to kill himself if Bush took up the presidency and Bush, knowing this, did so anyway and the threat was carried out, common sense does not impute the death to Bush but to the person who foolishly made and carried out the threat. Kant seems to agree, for he argues (in a far less convincing case) that even if one's refusal to lie to a murderer causally contributes to the murderer's finding one's friend and killing him, the death of the friend is imputed entirely to the killer, not to oneself for refusing to tell a lie. This suggests that, in a Kantian ethics, one must at least take into account whether deaths are a direct result of one's own actions (or inactions) or whether they result from someone else's immoral responses to what one does (or refuses to do). Judgments of responsibility after a death has been caused often take this factor into account, and so it seems also relevant as we deliberate before acting about what would be the right thing to do.

To expand our previous case, suppose that we have three options: Submit to terrorist demands, attack the terrorists, or refuse their demands and do nothing further. If we submit, let us suppose, we will save the hostages but almost certainly encourage further terrorism that will cost even more lives. If we attack, we will kill some hostages but will save more hostages. If we refuse to negotiate and do nothing else, the terrorists will probably kill all the hostages. In the first and third cases the deaths that occur will result through the agency of the terrorists themselves, even though our choice is a causal condition. In the second case (attack), the deaths will be a direct result of our actions (though we are responding to a problem posed by the terrorists). Prior to trying to figure out which course, all things considered, is morally best, we need to assess, in addition to our intentions and the estimates of lives expected to be lost and saved, whether it is also relevant whether the deaths are directly caused by us or by the terrorists.

Extending our previous line of thought, we could also address this question from the abstract point of view of Kantian legislators. The dignity principle itself should make us ex-

tremely reluctant to risk killing human beings, but it should also make us very hesitant to do something when we are confident that others will respond by killing human beings. Perhaps the refusal to take action that may kill others, other thing being equal, is more directly expressive of the way we value human lives; but other things are rarely equal, and refusal to take extraordinary steps to save lives can also express inadequate concern. We cannot be certain, of course, that the terrorists will carry out their threats if we do nothing; but then we are also uncertain how other options would turn out, and our question now is whether anything but these estimates of probabilities and numbers should make a difference. If it is generally worse to risk killing someone than to do what has the same probability of someone else doing a killing, then there must be reasons for this policy that are cogent from the Kantian legislative point of view. But what could they be?

The policy in question is a fairly specific application of a more general principle that is often maintained, namely, that it is generally better to "keep one's own hands clean" than to do what would normally be wrong in an effort to prevent wrongs that one anticipates that others will commit. This says, vaguely, that one's primary responsibility is to worry about the direct consequences of what one does or fails to do, without excessive regard to whether others will take the occasion to respond immorally.

Now I can imagine four considerations that might be offered in support of this principle for a world in which most are conscientious and even those who are prone to crime are often amenable to reason. First, a public policy that gives you greater responsibility for what you cause independently of the immoral responses of others might work to minimize the damage so often caused by well-meaning people in their efforts to prevent the wrongdoing of others. If most people are conscientious and many are bunglers, we would perhaps all do better if we do not encourage people to anticipate and act violently to prevent the crimes of others. Second, the policy would curb lawbreakers' power to get their way

through blackmail. If authorities were prepared to harm innocent persons whenever someone made a convincing threat to inflict more harm, then by using manipulative threats criminals could in effect make the authorities their accomplices. Third, the policy might promote the opportunity for most people to live out their own lives as rational agents because it places a higher priority on "keeping one's own house in order" than on intervening in the lives of those that we expect will act badly. By contrast, a policy that required everyone to be a moral busybody, always ready to do a lesser evil to prevent others from doing a slightly greater evil, would leave less room for planning and living out one's own preferred life. Fourth, the policy that makes each person primarily responsible for his or her own actions expresses a hope or faith that others will, after all, listen to reason despite their threats and past records; and so affirming that policy is a way of expressing respect for them as rational agents.

Though relevant, these policy considerations do not decisively support an inflexible "clean hands" principle, even in good conditions; and they are even less persuasive when violent crimes become more frequent and criminals more oblivious to reason. Moreover, the reasons that favor a "clean hands" policy for ordinary citizens do not apply equally well to law enforcement authorities. There are fewer risks and interferences with rational life plans when responsibility to prevent crimes is restricted to selected officials who voluntarily take up the task.

Furthermore, I suspect that the "clean hands" principle gains much of its popular support from a confusion. Suppose it is assumed that a certain act, say, telling a lie, is always wrong. Then if someone is tempted to lie to prevent another person from doing something worse, it might be appropriate to admonish the would-be liar by saying, "No, forget what others may do; your responsibility is for what you do." Similarly, if we assume that any intentional killing of a human being is wrong in itself, then even when killing one person would prevent others from killing many, we may say, "Keep your own hands clean; what others do is not your respon-

sibility." But the assumption in both cases is crucial. Unless we take as given that what we propose to do is wrong in itself, it remains an open question whether "what others do" affects our responsibility. And from the Kantian legislative point of view, what specific acts, if any, are always wrong and the degree to which we should take responsibility for preventing the consequences of others' crimes are matters about which we need to reason, not principles we can assume from the start.

CONCLUSION

My conclusions are limited and tentative. First, it seems that we can maintain the dignity principle without insisting that it is always wrong to kill human beings and to risk killing innocent human beings in response to terrorism. Second, there is a Kantian way of thinking about hard choices that admits, without abandoning the dignity principle, that at some level one must take into account the relative numbers of people whose lives can be prolonged by different policies. Third, though further discussion is needed, the objection that Kantian legislators would authorize some deliberate sacrifices of innocent persons seems dubious. Finally, although this Kantian perspective may give some support to a limited "clean hands" policy for most people under good conditions, it does not endorse the absolute principle "Regardless of what others threaten to do, your only responsibility is to ensure that you don't kill any innocent person."

Can the substantive principles "Don't kill human beings" and "Avoid endangering the lives of innocent persons" be maintained as principles in a Kantian morality even though they admittedly have exceptions? This depends on what "maintaining them as principles" means. Clearly not, if it means trying to keep the majority of people thinking of them as exceptionless rules while privately admitting that there are justified exceptions. But there may still be a sense in which the principles, with exceptions implicitly understood, might be held as quite fundamental principles, more so than

others (like "Don't commit incest") that have far fewer justified exceptions. That is, though qualified, the principles are partial expressions, in action terms, of the very attitude that the dignity principle itself demands, and the exceptions they contain are not concessions to a radically different moral point of view but rather refinements consistent with the basic Kantian viewpoint.

NOTES

This essay was written for the conference on violence and terrorism sponsored by the philosophy department at Bowling Green State University in November 1988, and subsequently discussed with faculty and students at the U.S. Airforce Academy, the College of Charleston, the College of William and Mary, UCLA, and the North Carolina Triangle Ethics Group. I want to thank the participants at these meetings for their many helpful comments.

- 1 Here I rephrase an often-repeated dialogue, sometimes attributed to G. B. Shaw, because the original version exploits sexist stereotypes and the differences are irrelevant to my point.
- 2 For simplicity I am imagining that my "critics" here are "act-consequentialists" and not "rule-consequentialists," for the latter might urge us in practice to treat certain rules as inflexible principles and yet defend this at a "higher level" of deliberation by appeal to consequences. The Kantian alternative to be considered, however, is opposed to both forms of consequentialism because of the constraints it places on the role of consequences in justifying principles.
- 3 Some of these points are spelled out in more detail in my "Humanity as an End in Itself," *Ethics* 91 (October 1980): 84-90, and (in slightly modified form) in my "Kantian Constructivism in Ethics," *Ethics* 99 (July 1989): 752-70. Many of Kant's works are relevant, but especially his *Groundwork of the Metaphysics of Morals*, trans. H. J. Paton (New York: Harper and Row, 1964), pp. 89, 97.
- 4 See Kant's *Metaphysical Principles of Virtue: Part II of the Metaphysics of Morals*, trans. James Ellington (Indianapolis: Bobbs-Merrill, 1964), p. 54.
- 5 Because of the centrality of the idea of "dignity," I shall here-

after refer to my reconstruction of Kant's principle that we should always treat humanity as an end in itself simply as "the dignity principle."

- 6 Kant, *Groundwork*, pp. 89, 97.
- 7 See Kant, *The Metaphysical Principles of Virtue*, p. 84.
- 8 Ibid., pp. 88-90, 96-100, 132-3.
- 9 Ibid., pp. 112-19.
- 10 Refusal to reduce moral decisions to weighing of costs and benefits does not, of course, imply that such weighing never has a legitimate role in moral thinking, even if some of Kant's more extreme remarks might suggest that he thought so.
- 11 Kant mentions punishment in several works but primarily in the first part of *The Metaphysics of Morals*, translated as *The Metaphysical Elements of Justice* by John Ladd (Indianapolis: Bobbs-Merrill, 1965). My view of Kant's position is presented briefly in "Kant's Anti-Moralistic Strain," *Theoria* 44 (1978): 131-51. For another view see Edmund L. Pincoffs, *The Rationale of Legal Punishment* (New York: Humanities, 1966), chap. 1.
- 12 This is a crucial move in the Kantian line of argument I am suggesting. It is similar to Rawls's procedural commitment to first fixing the moral constraints on the basic structure of society and only then working out principles for interactions between individuals. The idea is to take the dignity principle, as initially presented, to be inflexible but abstract, with its mode of application not yet specified. The examples in the *Groundwork* illustrate its application to interpersonal relations in a presupposed background of an ongoing society with a legitimate legal order. *The Metaphysical Elements of Justice* raises prior questions about the moral justification and limits of any system of public laws, with the assumption that day-to-day relations among individuals must conform to whatever justice demands regarding this necessary background. Thus the suggestion under consideration is that the dignity principle applies first to decisions about the basic system of public laws and only then to individual decisions remaining undetermined by this system. Though the priority implicit here may be doubted, it is not the inconsistent view that the absolute dignity principle applies independently to laws and to individual conduct and that when these applications conflict we should make exceptions in the individual sphere.
- 13 I draw here on a reconstruction of the Kantian legislative per-

- spective that I sketch more fully in "The Kingdom of Ends," *Proceedings of the Third International Kant Congress*, ed. Lewis White Beck (Dordrecht: Reidel, 1972), pp. 307-15. This view is expanded (and modified slightly) in my "Kantian Constructivism in Ethics." The basic idea is similar, of course, to the idea of the "original position" in John Rawls's theory of justice, but there are important differences.
- 14 See Kant's *Metaphysical Elements of Justice*, p. 35; and John Rawls's *Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971), chap. 1.
 - 15 Kant, *The Metaphysical Elements of Justice*, pp. 35-6.
 - 16 See *ibid.*, p. 101.
 - 17 It is worth reemphasizing here that the dignity principle does not say that prolonging the life of a rational agent is the supremely overriding value, though Kant's strong stand against suicide (with absence of cases of justified killing) in the *Groundwork* often suggests this conclusion to readers. The most obvious way to show that one treasures a physical object, say a "priceless" vase, is to do all one can to prevent its destruction, (almost) no matter what. But "rational agents" are not things like vases, and what it means to treasure them for what they are is accordingly more complex. Placing a supreme value on rational agents as such requires us not only to seek to preserve them but also to abide by the structures that they, as rational agents, endorse.
 - 18 These assumptions, of course, may not be satisfied in actual cases. Terrorists often claim high-minded motives and deny the justice of the legal systems they attack. Because of cultural differences we often cannot assume that the terrorists knowingly violate a shared moral framework, and in some cases they may not satisfy even the minimal rationality condition. To the extent that the assumptions fail, the line of justification we have been considering does not apply in any straightforward way.
 - 19 Because of his rigorist stand on certain issues (e.g., lying), contemporary discussions often simply label Kant an "absolutist" while treating any theory that allows that "the numbers count" as "consequentialist"; but this is misleading. It matters how, for what reasons, and under what constraints a theory allows the numbers to count. An interpretation (or extension) of Kant's theory that admits that sometimes one should take

- into account, at some level of deliberation, whether one or a thousand will be killed does not thereby "reduce" to consequentialism.
- 20 These other constraints are potentially important. My argument is not that any policy is acceptable if it would be approved by rational impartial legislators focused exclusively on survival. There is more involved in valuing human dignity than caring about prolonging the lives of rational agents. The point is that, unless circumstances reveal ways that a life-maximizing policy violates or undermines some other value implicit in human dignity, the dignity principle will allow and even encourage us to take into account whether the policy is likely to save the most lives.
 - 21 To focus discussion, assume that the victim in question is either unable or unwilling to volunteer and that he or she is law-abiding and is not among those at serious risk of being killed by the terrorist. Further, from now on I shall concentrate on cases where the sacrificial killing in question would be intended, not merely a "foreseen unintended consequence." (Killing a chosen victim to placate the terrorist is a clear example; shooting the terrorist *through* the hostage is more controversial.). Also, as before, I am restricting my remarks to official (e.g., police) responses to terrorism and assuming the victims are "innocent" in an intuitive, but still undefined, sense. In a fuller discussion, of course, other variations would need to be considered.
 - 22 These worries have been often and well expressed at the meetings where I have presented this essay and by an anonymous reader for the press. To these I owe not only thanks but, in time, a more fully developed response.
 - 23 It is worth keeping in mind here, as Larry Becker has reminded me, that any theory that attempts to resolve these dilemmas will leave conscientious persons unhappy because, despite the best of will, some things about which good people care deeply will be lost and our choices affect which these will be.
 - 24 The aim of the legislative thinking is to find public policies that officials should scrupulously carry out. This does not mean, of course, that in their deliberations the Kantian legislators should foolishly assume that in fact there will never be abuse and corruption by public officials.
 - 25 Treating someone as more than a "mere means," in the or-

- dinary sense, does not suffice to show that one regards that person as an "end" in Kant's sense. One might, for example, regard a cherished old car as more than a "mere means" without attributing to it an "unconditional and incomparable worth."
- 26 This was suggested to me by Andrews Reath. The point here is not that it is logically impossible to attribute an incomparable or incalculable worth to a person while one deliberately kills that person. Perhaps some community of superhuman beings, with complete self-control and purity of motive, could mutually agree on a policy of deliberate sacrifice when it would save more lives and then carry out the policy from pure devotion to duty, never once seeing this as an exchange of less value for more. But we are not like that. So, given human nature, it seems unlikely that we would, or even could, agree on and implement the policy "Deliberately kill innocent persons when it will save more" without coming to regard persons as things with a value subject to exchange.