Deontology at the Threshold

LARRY ALEXANDER*

TABLE OF CONTENTS

I. INTRODUCTION

In his 1989 law review article, *Torture and the Balance of Evils,*[1] later republished as Chapter Seventeen in *Placing Blame,*[2] Michael Moore declares himself to be a "threshold deontologist."[3] What he means is this: There are some acts that are morally wrong despite producing a net positive balance of consequences; but if the positive balance of consequences becomes sufficiently great—especially if it does so by averting horrible consequences as opposed to merely making people quite well off—then one is morally permitted, and perhaps required, to engage in those acts that are otherwise morally prohibited. Thus, one may not kill or torture an innocent person in order to save two or three other innocent people from death or torture—even though purely consequentialist considerations might dictate otherwise. However, if the number of innocent people who can be saved from death or torture gets sufficiently large, then what was morally proscribed—the killing or torture of an innocent person—becomes morally permissible or mandatory. At a certain number of lives at risk—the Threshold—consequentialist moral principles override deontological ones. Says Moore:

> It just is not true that one should allow a nuclear war rather than killing or torturing an innocent person. It is not even true that one should allow the destruction of a sizable city by a terrorist nuclear device rather than kill or torture an innocent person. To prevent such extraordinary harms extreme actions seem to me to be justified.[4]

Does allowing consequentialist concerns to override deontological prohibitions at the Threshold collapse deontology into consequentialism? Moore denies this.

> A consequentialist is committed by her moral theory to saying that torture of one person is justified whenever it is necessary to prevent the torture of two or more. The [deontological] view, even as here modified, is not committed to this proposition. To justify torturing one innocent person requires that there be horrendous consequences attached to not torturing that person—the destruction of an entire city, or, perhaps, of a lifeboat or building full of people. On this view, in other words, there is a very high threshold of bad consequences that

---

1. Michael S. Moore, *Torture and the Balance of Evils,* 23 ISRAEL L. REV. 280 (1989).
2. *See* MICHAEL MOORE, PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW 669 (1997).
3. MOORE, *supra* note 2, at 723.
4. *Id.* at 719.

must be threatened before something as awful as torturing an innocent person can be justified.[5]

Notwithstanding, Moore does admit that there is a *psychological* danger that people, once told that deontological prohibitions give way to consequentialist considerations at some point, will be much too quick to allow those consequentialist considerations to dominate.[6] After all, if one can kill or torture to prevent the killing or torturing of *N*, it is hard to see why one should not kill or torture to prevent the killing or torturing of *N*–1, or *N*–2, and so on, right down to the point where the number saved is only slightly larger than the number harmed. However, this psychological point, which may suggest that people should not be told that consequentialist considerations can ever override deontological prohibitions, in no way undermines the moral truth that at some point, deontological prohibitions do give way.

Moore acknowledges that threshold deontology might appear arbitrary and irrational. As he puts it, "Why should goodness of consequences not count at all and then, at some point, count enormously in the sense that it fully determines the rightness of action?"[7] Moore's answer to this question is that consequences always count, even below the threshold, but until the threshold is reached, consequentialist principles are outweighed by deontological ones. He analogizes the deontological norms to a dam, and the consequentialist considerations to water building up behind it. Eventually, if enough water builds up, it will reach and exceed the dam's height—which is analogous to the threshold of threshold deontology. "There is nothing arbitrary about thinking *both* that there is no spillover until the threshold of the dam's height is reached, and that each bit of water always counts in determining whether water will spill over the dam or not."[8]

Finally, Moore attempts to deflect the charge that any number for the Threshold (other than the number pure consequentialism would select) will be arbitrary. May we torture an innocent to save 500, or will 450 do? And if 450, why not 449? Moore replies that apparent arbitrariness cannot be the basis of any powerful objection to threshold deontology because:

---

5. *Id.* at 721-22.
6. *See id.* at 722-23.
7. *Id.* at 723.
8. *Id.* Moore attributes this analogy to Joseph Raz.

this is no more than the medieval worry of how many stones make a heap. Our uncertainty whether it takes three, or four, or five, etc., does not justify us in thinking that there are no such things as heaps. Similarly, preventing the torture of two innocents does not justify my torturing one, but destruction of an entire city does.[9]

Moore is surely not alone in holding that threshold deontology is the correct description of our moral reality. Thomas Nagel, in his essay *War and Massacre*,[10] put forward a similar position, though one that smacks more of a moral dilemma than as a moral guide.

> While not every conflict between absolutism and utilitarianism creates an insoluble dilemma, and while it seems to me certainly right to adhere to absolutist restrictions unless the utilitarian considerations favoring violation are overpoweringly weighty and extremely certain—nevertheless, when that special condition is met, it may become impossible to adhere to an absolutist position.[11]

> There may exist principles, not yet codified, which would enable us to resolve such dilemmas. But then again there may not. We must face the pessimistic alternative that these two forms of moral intuition are not capable of being brought together into a single, coherent moral system, and that the world can present us with situations in which there is no ... moral course for a man to take ....[12]

Even Robert Nozick, perhaps the most significant modern advocate of deontological (cum libertarian) side constraints, concedes the possibility of a threshold at which consequentialist considerations could override deontological prohibitions.[13] The question of whether these side constraints are absolute, or whether they may be violated in order to avoid catastrophic moral horror, and if the latter, what the resulting structure might look like, is one I hope largely to avoid.

It is my aim here to raise some doubts about the threshold deontologist's picture of morality. I am going to do so, not by a direct frontal attack on its plausibility, but by assuming at first that it is correct and then doing what Moore, Nagel, and Nozick (the latter self-consciously) eschew—namely, examining threshold deontology's structure in some detail. That structure turns out to produce a number of anomalous results. It seems, for example, more plausible to reject the existence of deontological thresholds and to assume instead either that deontological side constraints are absolute, or, as Nagel has hinted, that morality is thoroughly dilemmatic. I shall reach no strong conclusions on any of these points, however, since my principal aim is not to persuade you of any position other than that the structure of threshold

---

9. *Id.* at 724.
10. Thomas Nagel, *War and Massacre, in* MORTAL QUESTIONS 53 (1979).
11. *Id.* at 56.
12. *Id.* at 73-74.
13. *See* ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 30 n.* (1974).

deontology merits serious attention—attention that it has not heretofore received.

II. THE STRUCTURE OF THRESHOLD DEONTOLOGY

A. *Putting Aside the Aggregation and Identification Problems*

One issue that I wish to put aside, at least for the moment, is whether there are such things as moral catastrophes. Or, put differently, is it really a worse thing, morally speaking, for 500 to die than for only 1 to die? John Taurek has led me to believe that this is a serious question.[14] Even though it seems fairly obvious that it is better to save A's life than to relieve B's headache (even though the death and the headache are not choices for a single person who will suffer them, but will be suffered by separate people), it is less obvious that it is better to save the lives of A, B, and C than to save the life of D. For although the deaths of A, B, and C total three times as many deaths as the death of D, there is no one for whom his death is three times as bad as anyone else's. Each person's death is equally bad for him.

Taurek raises the question whether similar harms to different people should be aggregated in considering where our moral duties lie. Most philosophers who have considered Taurek's argument have resisted it and have argued in favor of aggregation, and I shall assume, with reservations, that they are right and Taurek is wrong.[15] If Taurek were right, however, then threshold deontology would be an impossibility. If it is morally wrong to kill or torture one person to save two lives, then it would be morally wrong to do so to save N lives, no matter how large the N. For according to Taurek, the loss of N lives is not worse than the loss of two lives, no matter how large the N.

I shall also put aside for the moment the problem of identifying the threshold. Moore acknowledges that this is a practical problem, but he denies that it is a theoretical one, analogizing it to the sorites problem of determining when there are enough rocks for a heap, or few enough hairs for baldness. I believe, to the contrary, that underlying the practical problem is a theoretical one, and one that is different from the sorites problem of vague predicates such as baldness. But I shall

---

14. *See* John Taurek, *Should the Numbers Count?*, 6 PHIL. & PUB. AFF. 293, 293 (1977).

15. *See, e.g.,* 1 F.M. KAMM, MORALITY, MORTALITY: DEATH AND WHOM TO SAVE FROM IT 75-98 (1993).

approach this point obliquely by assuming for the time being that we can identify the precise numbers for the thresholds of threshold deontology. (I use plurals rather than singulars in referring to numbers and thresholds because it may be more plausible to assume that there are different thresholds depending upon what harms are to be inflicted and what harms are to be prevented thereby.)

1. The Isomorphism of Negative and Affirmative Duties with Respect to Thresholds

I believe that the most plausible account of a deontological threshold would consider both the number of persons at which the interests at stake justify overriding negative duties (such as those forbidding killing, torturing, and removing blood and organs from the non-consenting) and the number of persons at which the interests at stake justify imposing affirmative duties to rescue, give up one's organs, and perhaps sacrifice one's life. For example, suppose the death of X will save N lives. If N is the number of lives at stake that serves as the threshold deontologist's point at which it becomes permissible to kill X, then, of course, it follows by definition that it is justifiable to kill X to save N. Suppose, however, that we are unable to kill X, but that X could kill himself and thereby save N. Although X generally has no affirmative duty to sacrifice himself to save others (by giving up his life, giving up a kidney, or even by making the slight effort and taking the slight risk of an easy rescue), it is also true that he generally has a right, corresponding to others' negative duties, not to be killed or otherwise treated as a resource to save others without his consent. If the latter right is overridden at the threshold, then it is difficult to see why this would not entail the overriding of the absence of an affirmative duty to save, and thus, the creation of an affirmative duty to save. For if I can kill X to save N, then I should be able to threaten X with harm if he does not attempt to save N, even at a high risk to himself—indeed, even if the rescue involves certain death for X. And if I can justifiably threaten X with harm unless he rescues N, it must be because imposing an affirmative duty on X to rescue N is justifiable. And imposing such an affirmative duty would be justifiable if there indeed exists such an affirmative duty where the deontologist's threshold is reached.

2. Is the Threshold a Ratio?

There are two ways one can conceptualize the deontologist's threshold. One way is to conceptualize it as a point at which, for every life at stake beyond the threshold, another life can be taken. Thus, if one

life can be taken to save N lives, two lives can be taken to save $N+1$ lives, three lives for $N+2$, and so on. (The same would be true for the thresholds for lesser interests: blood can be taken from one person to save a specific number of lives, from two people to save one more than that number, and so on.)

The alternative conceptualization of thresholds is to see them as ratios expressing the relation between the number of persons to be saved and the number of lives or other interests that can be sacrificed to save them. Under this conceptualization, one may kill one person to save N, two people to kill $2N$, three people to kill $3N$, and so on.

Moore's metaphor for the deontological threshold—that of a dam holding back a lake—seems to suggest the first conceptualization, on which each additional life (or other interest) at stake beyond the threshold justifies the taking of an additional life (or other interest). For once the level of the dam is reached, one needs a foot of additional sandbagging to hold back each foot of water above that level.

However, notwithstanding the dam metaphor, the conceptualization of ratios seems more plausible. First, it explains why we might think that the threshold might be lower if the negative right we override is not the right not to be killed but, say, the right not to have a pint of blood taken without consent, or the right not to be touched without consent. One may not be permitted to kill X to save $N-1$ lives, but one should be able to touch him or take his blood to save some number of lives less than N. Indeed, assuming for the sake of argument that we are still threshold deontologists about these matters, one should be able to touch X without his consent to save one life, or even to save someone else from considerable pain. These intuitions suggest that the deontological threshold might be some ratio between (1) the interests at stake times the number of people affected in the to-be-saved column and (2) the interests at stake times the number of people affected in the negative-rights-overridden column.

The second point in favor of the ratio conceptualization of the threshold is this. Suppose that for killing 1 innocent person in order to save other innocent people from death, the threshold is 100. That is, only when 100 lives can be saved is it morally permissible to kill 1 innocent person. Now suppose that A, B, and C are in the hospital for minor surgical procedures. Suppose further that each of A's, B's, and C's tissue, organs, and so forth could be used to save 99 other people from death. By hypothesis it is not permissible to kill A to save 99, B to

save 99, or C to save 99. Under the first conceptualization of the threshold (that views N as the point beyond which one can take 1 additional life to save an additional life) one can kill A, B, and C to save 297 people. But that conclusion conflicts with the impermissibility of killing each to save 99.

The same point can be made slightly differently. If the threshold is merely the point beyond which lives can be taken to save an equal number of lives, then what stops us from saying that given that far more than N lives have already been lost—as is surely the case—we should from this point forward act solely as consequentialists? I can see no good answer to this question; thus, I conclude that any plausible threshold deontology will involve ratios, probably not only involving the numbers at stake, but also involving the interests at stake.

III. APPLICATIONS

If there is a deontological threshold or thresholds at which negative rights give way to affirmative duties, then one might expect that moral outcomes in the neighborhood of such a radical discontinuity might seem weird or counterintuitive. And that expectation is borne out. In the examples that follow, let us make these assumptions:

(1) The threshold (N) at which one can *torture* an innocent person to effect a saving of *lives* is 100 lives at stake. Put in shorthand, $N_{torture/lives} = 100$.

(2) The threshold at which one has an affirmative duty to undertake a minimally difficult and minimally risky rescue is 20 lives at stake; i.e., $N_{easy\ rescue/lives} = 20$.

A. Saving Fewer Than N

1. The Terrorist's Mother

A terrorist has captured and hidden 100 hostages in a number of locations. At each location he has planted a bomb that will automatically detonate at noon. The terrorist reveals this to the police and demands certain concessions from the government before noon; otherwise, he will allow the bombs to kill the hostages.

Luckily, the police have grabbed the terrorist's mother, who is in no way implicated in her son's terrorism. If they torture her, however, they believe that the terrorist will relent. And because 100 lives are at stake, the mother's right not to be tortured is overridden. So they begin torturing her on live television. Sure enough, after some severe suffering by his mother, the terrorist calls the police asking them to relent. The police ask him for the location of the hostages and the bombs, and he tells them where one, and only one, hostage and bomb are located. Because only 99 hostages now remain at risk, the police can no

longer torture the terrorist's mother.

2. *The Rescue*

> Twenty people are clinging to a housetop surrounded by rapidly rising floodwaters. Rescuer, who has a large boat that can easily and safely hold the threatened twenty, understands himself to have an affirmative duty to rescue them. As he pulls alongside the housetop, one of the stranded people jumps off and into Rescuer's boat. Rescuer, realizing that there are now only nineteen people in harm's way, turns away from the housetop and ignores the desperate pleas of the nineteen left behind.

I can see no possible response by the threshold deontologist to the last two cases except to concede that the torture of the terrorist's mother must cease once the single hostage can be saved, and that the rescuer violates no duty if he takes the single flood victim from the housetop and leaves the remaining ones to die. What should be unsettling to the threshold deontologist, however, are not these conclusions per se, but that these conclusions suggest that the other lives at risk are pure moral ballast, serving no purpose other than to create a duty to save someone else, or to override a right that stands as an obstacle to saving someone else. Like the water below the level of the dam, their only moral role is to boost others above that level so that they become morally entitled to be rescued.

Moreover, looked at from the standpoint of Terrorist's mother or Rescuer, cases like these that are right at the threshold are really cases where negative rights are overridden and affirmative duties are created merely to save one person's life (the one hostage whose whereabouts are revealed, or the one flood victim who jumps into the boat). For once the single life is saved, Terrorist's mother has her right not to be tortured restored, and Rescuer is relieved of his affirmative duty to rescue. By hypothesis, however, one may not be tortured or placed under an affirmative duty to rescue for the sake of one life. So cases at, or just above, the threshold seem paradoxical: the very deontological constraints that protect one from straight consequentialist trade-offs below the threshold disappear above it, so that one then can be killed, tortured, or forced to rescue merely to save one person.

The same point holds for killings in order to save lives. For suppose we can kill A and save N lives, or B and save $N-1$ lives. If we must choose A and not B to kill, from A's standpoint he is being targeted for death for the sake of one life.

B. Increasing the Danger to Reach N

1. Terrorist's Mother II: The Luring

Suppose this time that the terrorist has hidden only one bomb and ninety-nine hostages. Because $N_{torture/lives}$ is 100, the police are stymied. However, they do know approximately where the bomb and hostages are located, and they can, through some deception, lure a handicapped man in a wheelchair into that area. Presently, he is far enough from the bomb to escape death, though there is some risk that he will be injured when it explodes. If lured close to the bomb, he will be unable to wheel himself out quickly enough to escape death from the noon blast. The police proceed to lure him into harm's way, at which point they begin torturing the terrorist's mother, with the result that the terrorist reveals the location of the bomb and hostages.

2. Rescue II: The Luring

This time, there are only nineteen flood victims on the housetop. Rescuer will not rescue them. If, however, he has an affirmative duty to do so (if there are twenty people at risk) then he will discharge that duty and will rescue as many as his boat can hold. The nineteen see one person rowing a tiny rowboat, which is in some (slight) danger of capsizing, and which is too far from any place above water for the rower to escape drowning if his boat does capsize. The nineteen lure the rower near their refuge, at which point they pummel his rowboat with rocks, causing it to capsize and him to swim to the housetop. Because there are now twenty people at risk, Rescuer now has an affirmative duty to rescue them, which he does.

In both these cases, it may appear that the rights of the handicapped man and the rower have been violated. Yet, by being lured into a situation in which they faced certain death, their welfare was arguably increased. Before the luring, they faced risks of injury and drowning. After the luring, they faced tiny risks of death, since the police were confident the torture would work, and the flood victims were confident that Rescuer would do his duty. Had the handicapped man been asked to go near the hostages, or the rower asked to destroy the rowboat and join the others on the housetop, each might have done so voluntarily had he expected the torture or rescue to then commence. But the threshold deontologist would have reason to reject voluntarily created thresholds as valid ones—that is, as thresholds that overcome negative rights and create affirmative duties. After all, think of the incentives that would be created if we could override deontological constraints by voluntarily becoming part of N. Deception might thus have been the only method for increasing the welfare of the men in the wheelchair and rowboat.

So the question is, is it wrong to deceive or coerce people in order to increase their welfare along with others' when they cannot voluntarily do so themselves? And whatever the answer to that question is

generally, is it wrong to do so in order to be able to override others' deontological rights? Or is it wrong to do so because one cannot then override the deontological rights that, if overridden, would make the deception or coercion welfare-increasing?

The luring cases are really examples of a more general problem. Suppose that the deontologist's threshold for overriding the right not to be killed is the saving of 200 lives. In other words, $N_{killing/lives} = 200$. Now consider the following case.

3. Surgeon

> Surgeon has a patient under anesthesia for a minor procedure. There are 199 other patients in the hospital who are in imminent danger of dying unless they receive the anesthetized patient's rare blood. Unfortunately, taking that much blood from the patient would kill him. Surgeon has another patient in the hospital for a condition that is debilitating but not life-threatening. That condition, too, can be treated by transfusing the first patient's blood. Surgeon can worsen this last patient's condition to the point that it becomes imminently life threatening. He does so, without the patient's consent or knowledge, then kills the first patient and saves the 200 dying ones.

Surgeon is like the "luring" cases described previously, in that the case involves an act that is a prima facie violation of a deontological right—namely, the intentional worsening of the last patient's condition. And like the luring cases, the result of this prima facie violation is an increase in the welfare of the one whose rights are prima facie violated.

What this entire class of cases suggests is that a deontological threshold invites manipulation in order to produce consequentialist-justified results. And if the threshold deontologist rejects the moral possibility of such manipulations, he owes us an account of why they are wrong. Specifically, are they wrong because they violate the rights of the one deceived, coerced, or physically manipulated into the dangerous condition? And if so, why are the manipulations rights-violative, given that they are welfare-increasing and that the rights-holders would gladly consent to their endangerment if consent were possible without morally precluding the acts that remove their endangerment?

Or are these manipulations violations of the rights of the ones who are ultimately sacrificed (killed, tortured, or forced to rescue)? If so, what explains this, given that in every case, they are sacrificed to save N, and none of the N has voluntarily endangered himself?

C. How Many Violations?

1. Failure of Easy Rescue

Recall that $N_{easy\,rescue/lives} = 20$. Suppose that there are exactly twenty people on the housetop in the flood. By hypothesis, Rescuer has a duty to rescue them. Suppose he refuses to do so, and they all drown. Is he guilty of twenty murders (by omission) or only one? There are twenty corpses; but had there been one fewer, Rescuer would have been entirely off the hook. And as we saw in "Easy Rescue," he would have been off the hook had he rescued only one of the twenty. But if he is guilty of violating the rights of only one victim, which victims' heirs can sue him for wrongful death?

2. Terrorist's Siblings

Again, Terrorist has hidden a bomb and 100 hostages, and $N_{torture/lives} = 100$. This time, the police have captured Terrorist's innocent brother and sister. Because they are unsure whether Terrorist will break if only one of his siblings is tortured, they torture both until Terrorist reveals the location of the bomb and hostages.

By hypothesis, the police have acted immorally. One person may be tortured to save 100, but not two. But are they guilty of violating the rights of *both,* or only one, of the siblings? If the latter, whose?

These cases illustrate (from the perspective of both those who are in danger and those who are used as means to avert the danger) that no matter how many people are harmed when people act immorally in a threshold situation, there is always some alternative course of conduct that is morally permissible in which almost as many people are harmed. Although the immoral *conduct* (or omission) causes great harm, its *immorality* causes much less. That is because threshold deontology treats those endangered as moral ballast up to the threshold point, and it treats those whose sacrifice can save the endangered as expendable resources once the threshold is reached.

D. Probabilities

In all of the examples thus far, we have regarded the deaths of the potential victims as certain. But certainty about such matters is rare. If $N_{killing/lives} = 200$ when we are certain that the 200 will die, what is it when we believe that killing has only a 50% chance of saving lives? Is N 400 in such a case? Is it 400 when we are certain that killing will save the lives, but only 50% certain that the 400 will die without the killing? Do the probabilities work this way, such that as we become more uncertain about the efficacy of killing or the danger if we do not kill, N increases in the same ratio, so that a killing with only a 1% chance of saving the

lives is justifiable only if 20,000 lives are at stake? Or does N vary in a more complicated way? If the probability becomes sufficiently small, do the rights violations become impermissible no matter how large the N?

E. Time-Framing of Moral Catastrophes

Suppose Fred's liver has marvelous properties and that, were Fred's liver removed, killing Fred, many people's lives could be saved. $N_{killing/lives} = 200$, so that if only 199 lives are at stake, it is wrong to kill Fred for his liver. Now suppose that every month, 199 people are in need of the properties in Fred's liver. If they do not receive them, they will die. Fred's liver can save more than N. May we kill Fred, extract his liver, save the 199 one month, and then save some of those who become ill in subsequent months? Or is each month's new wave of illnesses a separate moral catastrophe, which cannot be aggregated with any other month's catastrophe for purposes of exceeding N? Is N time-bound (or space-bound), and if so, why? If not, may Fred be killed in order to save a person per year for the next 200 years? And, coupling this issue with the issue of probabilities raised in the previous subsection, may Fred be killed for, say, cancer research if we believe that by killing Fred—and only by killing Fred (or someone similar)—we have a $1/X$ probability of saving NX people from cancer deaths?

IV. THE ARBITRARINESS OF DEONTOLOGICAL THRESHOLDS

Up to this point, I have been identifying puzzles and raising questions about deontological thresholds without assuming either that aggregating harms is improper or that the numbers of the various Ns are themselves arbitrary. I shall continue to assume, in contrast to Taurek, that aggregation of harms is not a problem. Granted, there is more than a whiff of Taurek's point in those examples in which all of the lives at risk are merely so much moral ballast to justify the saving of the Nth life.

What I intend to do now is drop the assumption that N is not arbitrary in a morally problematic way. That assumption, I believe, is incorrect. Recall that Moore believes that locating N, while difficult, is no different in kind from the sorites problem of identifying the number of stones required to make a heap. I believe, on the contrary, that the moral arbitrariness of N is not a sorites problem, and that such arbitrariness seriously threatens Moore's picture of deontological thresholds.

My demonstration of *N*'s arbitrariness will borrow heavily from a recent article by Anthony Ellis.[16] Ellis puts forward three propositions that a threshold deontologist might accept:

(1) There is some number [*N*] at which the act that was wrong becomes right....
(2) There is no non-arbitrary way of specifying [*N*].
(3) The difference between what is morally right and what is morally wrong cannot be arbitrary.[17]

The three propositions are inconsistent, so one must be rejected. The threshold deontologist presumably must hold (1), so that leaves rejecting either (2) or (3). And (3) seems plausible, so (2) appears the most eligible for rejection.

Is (2) correct? Ellis canvasses several arguments that might be raised against (2). First, someone might assert that the threshold for killing to save lives is ten, and then defend that choice of *N* against the charge of arbitrariness by pointing out that the death of ten is worse than the death of nine. But, says Ellis, this cannot rebut the charge of arbitrariness. As Ellis puts it, "[W]hat needs to be explained is why a given level of harm generates the [threshold], and this is not explained by pointing out that this level is greater than some lesser level."[18]

Neither is it a satisfactory defense of the non-arbitrariness of *N* to appeal to the idea of *judgment*. Ellis's response to this tack, which response I endorse, is worth quoting at length:

> This response seems—to speak plainly—to be little more than bluff and rhetoric. There can be judgment only where there is something to be judged *about*. And that does not simply mean that there must be an issue to be resolved, but that there must be considerations capable of resolving the issue (or—at the very least—bearing intelligibly upon its resolution). *And the problem here is to see what considerations could do that.* It is not enough simply to assert that we have to weigh the wrongness of the action against the badness of the consequences of refraining from it. That is not the solution, but the problem. What we need an account of is how that weighing can be carried out—an account of what considerations judgment is supposed to be exercised upon, and how that exercise works upon them.[19]

Put simply, we need to be told not merely that locating *N* requires judgment, but how judgment is to take into account the relevant considerations of lives and rights.

A third response that Ellis rejects is one that argues that the problem of locating *N* is merely epistemological. For example, many people hold such a view about the morality of abortion. Such people believe that we

---

16. *See* Anthony Ellis, *Deontology, Incommensurability and the Arbitrary*, 52 PHIL. & PHENOMENOLOGICAL RES. 855 (1992).
17. *Id.* at 859.
18. *Id.* at 860.
19. *Id.*

know that at some point between conception and birth, the fetus becomes a "person" (i.e., morally protectable), but that we do not know exactly where this point is. Ellis replies that even if the fetus becomes a person at a precise point in time, and it does so by virtue of something that we cannot understand (e.g., ensoulment by God), the same thing cannot be true for locating *N.*

> The present suggestion is that there is a reason why [*N*] is just here but one which we do not know, or cannot understand. But the suggestion seems bizarre. It is hard to believe, after all, that there is anything more that we could learn about the scale of harm which [*N*] ... involves, and its relation to the harm involved in other [situations], which could even be relevant to our moral decision.[20]

Locating *N* is not an epistemological problem.

The reason locating *N* is arbitrary, Ellis argues, is that it involves the weighing of incommensurables.[21] Deontologists treat killing an innocent person for others' benefit as *intrinsically* wrong. Consequences are immaterial insofar as one is a deontologist. On the other hand, for the consequentialist, consequences are all that matter. The only thing that is intrinsically wrong for her is the act-type "failure to promote the best consequences."

Consequentialism and deontology may be incommensurable in two ways. One might say that both deontological norms and consequentialist ones apply to every choice, but neither is weightier than the other.[22] When they generate conflicting prescriptions, then they create moral dilemmas. But this will not help the threshold deontologist locate *N* because the threshold deontologist holds, not that there are moral dilemmas where consequentionalism and deontology conflict, but that deontology should reign below *N* and consequentialism above it. That is an entirely different picture from one of moral dilemma.

The other sense in which deontology and consequentialism might be incommensurable is if *no* consequentialist considerations could ever outweigh deontological ones (or vice versa).[23] But this would lead to regarding deontological (or consequentionalist) norms as *absolute.* The threshold deontologist obviously must reject this account as well, for on it, just as on the moral dilemma account, there is no *N.*

---

20. *Id.* at 861-62.
21. *See id.*
22. *See id.* at 862-63.
23. *See id.* at 863.

The threshold deontologist must regard deontological and consequentialist norms as being capable of being weighed against each other—that is, as being commensurable. For the threshold deontologist's claim is that at N, consequences *justify* acts that are otherwise intrinsically wrong. But it is difficult to see how consequences could do this unless consequences and intrinsic wrongness were commensurable on some scale.[24]

If the specification of N is arbitrary—and I believe that Ellis has shown that it must be—then perhaps the threshold deontologist should reject either the third proposition (that the difference between what is morally right and morally wrong cannot be arbitrary), or the first proposition (that there is a specific number N at which wrong acts become right) mentioned above.

I believe it is clear that the threshold deontologist—as opposed, perhaps, to someone who viewed morality as thoroughly dilemmatic—cannot accept that moral rightness is arbitrary and thus must accept the third proposition. That leaves the threshold deontologist with the possibility of rejecting the first proposition. Indeed, Moore himself regards the problem of specifying N for threshold deontology as analogous to the classic sorites problem of specifying N for rocks in a "heap."[25] Just as there may be no specific N for "heapness," there may be no specific N for killing, torturing, or rescuing, and so forth, in order to save lives, avoid pain, and so forth. As Ellis puts this argument, killing one to save one may be wrong, killing one to save two may be wrong (but less so), and so on, "[b]ut . . . as we proceed through the series of acts the act gets less wrong until, at some later points, it is right; but there was never any *point at which it became right*. There is in the series no last case of rightness."[26] Just as if we start with a totally bald man and add hairs to his head, eventually he would be hirsute without there being any specific number of hairs that marks the point where baldness ceases and hirsuteness begins, so too with the morality of killings to save lives.

Now this analogy requires that there be degrees of wrongness. In the typical sorites example, we have a vague predicate (e.g., "bald" or "orange") and a smooth continuum (e.g., from bald to hirsute, or from orange to red), and it does make sense there to refer to degrees of baldness or to degrees of orangeness. However, in the case of threshold deontology, there is a radical discontinuity at the point of N. So in order to make use of the sorites analogy, the threshold deontologist must reject

---

24. *See id.*
25. *See* MOORE, *supra* note 2, at 724.
26. Ellis, *supra* note 16, at 866.

the idea that there are discontinuities as we move from deontologically forbidden to consequentially required.

Suppose the threshold deontologist rejects discontinuous change and accepts that there are diminishing degrees of wrongness as we make the consequences of abiding by deontological norms more and more dire. Can he then, as in the case of vague predicates, avail himself of the sorites analogy and deny that there is a specific number N that marks the threshold? He cannot. Denying that there is a specific cutoff point in threshold deontology is arbitrary and troubling in a way that is disanalogous to typical sorites examples. It is worth quoting Ellis at length on this point:

> Imagine that someone held that act 56 was wrong but that act 57 was right, but that the change from wrong to right here occurred in a continuous way. *The claim that the transition occurs continuously makes it no less arbitrary to claim that it occurs between 56 and 57.* And nor would it be any less arbitrary to claim, what would sound more realistic, that it occurred roughly between 50 and 60, or between 7 and 20, or between any other numbers (including the 1 and 400, which, for the sake of the argument, we gave to the non-absolutist deontologist at the start). The postulation of degrees of wrongness, whatever intrinsic merit it may have, does nothing to solve *that* problem. So the denial that there is a precise cutoff point, which is what is at the heart of this response, is beside the point.
>
> Since the response is such a popular one, the point may bear repetition. Take an analogy. There is no precise cutoff between, say, red and orange, and if we wanted one we should have to specify it arbitrarily. But we could not put it just anywhere within the colour spectrum. We should be faced with specifying an arbitrary point, but within a non-arbitrary range. That is not what we are faced with in morality. Specifying a range where the transition from wrong to right takes place would be no less arbitrary than would be specifying a precise cutoff. And if, to make the analogy with the transition from red to orange yet closer, it was said that the limits of the range are somewhat vague, this would still leave the position unaltered. The concession that the range is a vague range makes it no less arbitrary.[27]

Why is the threshold deontologist's location of N disanalogous to the problem of vague predicates, such as red and orange?

> In the transition from red to orange we find a range of cases about which we are unsure (itself a vague range) and ranges of cases (vague again) about which we are quite sure. Now those ranges are given by a deep and wide agreement in judgements, and that agreement settles any questions that might arise about the proper description of any point on that part of the colour spectrum. It could not be like this in morality. It is not just that we do not find agreement in judgements here. (Though we don't, and this is to be expected on my account.) If that were all that it is, then we could simply conclude that, where there is not

---

27.   *Id.* at 869.

such agreement, the matter is indeterminate, not yet decided. The problem is that in morality such judgements.... cannot be settled by group agreement.... Suppose that everyone did in fact agree, on the number 50 say, *but no-one could give any reason why it should be 50 and not some other number.* This would not tell us anything about moral theory; it would simply be an utterly bizarre mystery. And if we found agreement on a vague range—again with no-one able to give any reason to justify this range rather than some other—then this might be less dramatic but it would be just as mysterious. It would be senseless to think that this agreement could make any contribution to settling the moral question.[28]

Thus, the threshold deontologist cannot utilize the sorites analogy and deny that there is any specific N that marks the threshold. And I have already shown that any specific N must be arbitrary. If one then accepts the third proposition and denies that the difference between the morally right and morally wrong can be arbitrary, threshold deontology would then be an untenable position. The three propositions are not jointly compatible, yet the threshold deontologist must accept them all.

V. THRESHOLD DEONTOLOGY AS INDIRECT CONSEQUENTIALISM

Perhaps, however, I am mistaken in thinking that threshold deontologists like Moore are really deontologists at all, even if a softer type than Kantian absolutists. Suppose that threshold deontologists are actually thoroughgoing consequentialists for whom deontological restrictions are at bottom merely heuristics for the indirect pursuit of the Good. This would explain Moore's rather cavalier treatment of how N is to be specified. Moore could believe that following rules that prohibit killings or torturings of some as means to save others from greater harm would ultimately produce better consequences than having each person weigh the predicted consequences of each choice. The reasons are familiar—cognitive limits and biases, coordination difficulties, and decision costs.[29] On the other hand, a moral rule which tells us not to kill or torture unless the consequences of not doing so are catastrophic, but that leaves catastrophic undefined, might produce better consequences than one that absolutely forbids killings and torture, or one that specifies the precise number of saved lives that justify killing or torture. The absolute rule might be followed too well. The rule with a specific threshold may look too arbitrary (because it *is* arbitrary), and people might start reasoning down the slippery slope from an arbitrary N to pure consequentialist balancing, which by hypothesis is undesirable.

A rule with an exception for catastrophes, but that leaves catastrophes

---

28. *Id.* at 869-70.
29. *See* LARRY ALEXANDER & EMILY SHERWIN, THE RULE OF RULES: RULES, PRINCIPLES, AND THE DILEMMAS OF LAW (forthcoming 2001).

910

unspecified, may then be optimal in terms of consequences. Of course, the matter is an empirical one and like all indirect consequentialist rules cannot be confidently formulated from one's armchair. Still, the account that sees threshold deontology as an indirect consequentialist strategy is plausible, and it can explain why the inability to specify N non-arbitrarily is not a problem but a virtue.

In the end, however, I doubt that Moore—or Nagel or Nozick for that matter—want to grasp this lifeline. Nothing in what Moore or the others have written suggests that they are consequentialists at the level of their deepest moral beliefs. To the contrary, I believe that their deontological positions are at least as fundamental are their consequentialist ones. If I am correct, then the arbitrariness of N cannot be dismissed by Moore in the way that an indirect consequentialist might.

VI. CONCEPTUALIZING THE CONFLICT BETWEEN DEONTOLOGY AND CONSEQUENTIALISM

In the end, I believe that Moore's picture of the conflict between deontological norms and consequentialist ones is erroneous. In my opinion, the best conception of deontology would deem its core principle to be that one may never use another as a resource without his consent.[30] In other words, a person's body, labor, and talents do not exist for others' benefit except to the extent that he freely chooses to benefit others.

Outside the realm of appropriating others' bodies, labor, and talents, deontology is inapplicable. Thus, in settling conflicts over uses of natural resources and incompatible activities, consequentialist considerations will be just as relevant to the deontologist as they are to the consequentialist. Although some deontologists see a role for a non-consequentialist "corrective justice" in this domain, I have serious reservations about extending deontology beyond purposeful appropriations of others' bodies, labor, and talents.[31]

Therefore, the deontologist can, like the consequentialist, consider consequences to be morally relevant *within the domain beyond the deontological core.* But he cannot, under any circumstances,

---

30. *See* Larry Alexander, *Is Morality Like the Tax Code?*, 95 MICH. L. REV. 1839, 1845-47 (1997).
31. Larry Alexander, *Causation and Corrective Justice: Does Tort Law Make Sense?*, 6 LAW & PHIL. 1, 7-11 (1987).

countenance the view that people are resources for others' benefit. It is his rejection of that view that defines him as a deontologist.

The consequentialist, on the other hand, does view each of us as resources for others. If my labor or my body can be used to maximize total welfare (for maximizing versions of consequentialism), or to equalize total welfare (for equalizing versions), then I am obligated to sacrifice, and others are permitted to take, my labor or my body. That we *are* resources for others is as fundamental to the consequentialist as that we are *not* resources for others is to the deontologist.

Now Moore's picture of threshold deontology as a deontological dam holding back a consequentialist body of water ultimately treats the deontological and consequentialist conceptions as commensurable. Dams and bodies of water can be compared along the dimension of height, so that the dam either will, or will not, be higher than the water. But deontology and consequentialism are incommensurable because they are fundamentally opposed conceptions of what morality is about. One sees the individual as inviolate, an end in himself, and the opposite of a resource for the betterment of the world. The other sees the individual in exactly the opposite way.

The threshold deontologist would have us believe that we switch from not being resources for others to being resources for others when N is reached. When N is looked at like that, however, it seems downright implausible that the moral universe is so constituted. There may be thresholds at which new phenomena emerge, but it is quite another thing to have thresholds at which things become their opposites.

912