

Chapter 9

Torture and Method in Moral Philosophy

Jeff McMahan

Introduction

Of the six philosophers who spoke at the conference at which a number of the papers in this book were presented, two devoted their talks to advancing objections to the hypothetical example known as the “ticking bomb case.”¹ In doing so, they joined a large chorus of distinguished contemporary moral and political philosophers who have criticized the use of that example.² The conference’s keynote speaker, Albie Sachs, went further by repeatedly criticizing all uses of hypothetical examples in philosophical discussions of torture. In the first part of this essay, I will offer some reflections on these protests against the use of the ticking bomb case and other hypothetical examples, or thought experiments, in debates about the morality of torture. In the second part, I will explore the ways in which certain distinctions in normative moral theory, the significance of which is usually demonstrated through the use of hypothetical examples, might be relevant to understanding the morality of torture.

1. Method in practical ethics

Many moral philosophers, particularly consequentialists but also others influenced by recent work in what is known as “experimental philosophy,” reject appeals to moral intuitions as a basis for moral theorizing and therefore tend to regard the use of hypothetical examples to elicit moral intuitions as an illegitimate method in moral theory. These philosophers argue that moral

¹ The talks by Marcia Baron (a revised version of which is published as Chapter 8 of this volume) and Claudia Card.

² See, for example, David Luban, “Liberalism, Torture, and the Ticking Bomb,” *Virginia Law Review* 91 (2005): 1425-61; Henry Shue, “Torture in Dreamland: Disposing of the Ticking Bomb,” *Case Western Reserve Journal of International Law* 37 (2006): 231-39; David Luban, “Unthinking the Ticking Bomb,” in *Global Basic Rights*, eds. Charles R. Beitz and Robert E. Goodin (New York: Oxford University Press, 2009), 181-206; and Claudia Card, “Ticking Bombs and Interrogations,” *Criminal Law and Philosophy* 2 (2008): 1-15.

intuitions in general, whether about hypothetical cases or actual cases, are epistemically unreliable; they are the products of our evolutionary heritage, adapted to the primitive conditions in which our remote ancestors lived, and have also been shaped by cultural prejudice, self-interest, and a variety of other distorting factors.

Most of the prominent critics of the ticking bomb case and other hypothetical examples in the debate about torture are not, however, among the skeptics about moral intuition in general. They do not object to appeals to moral intuition, but only to the use of certain examples, such as the ticking bomb case, or to the use of hypothetical examples as opposed to actual, historical examples. For example, in his contribution to this volume, which is a revised version of his keynote address at the conference from which this book evolved, Albie Sachs repeatedly dismisses appeals to hypothetical examples and claims that in societies in which people still bear the memory of their own government's use of torture against them, there cannot be debates about the morality of torture in which the participants "convert what is a huge historical experience still haunting the soul of the nation, into a set of imaginary situations to be dealt with at a purely instrumental level, weighing up profits and losses in an analytical book-keeping manner."³ Thus, as a citizen of such a society, he told the other conference participants who attended his keynote address that he could not attend their sessions, as it would be "horrific" to discuss torture by "concocting imaginary situations."

But Sachs is not opposed to reliance on the evocation of moral sentiment or intuition in general. On the contrary, his address, which appears in this volume, is essentially a sustained appeal to moral sentiment. It is just that the examples he describes, drawn largely from his own experience, are events that actually occurred. His restrained descriptions of actual instances of torture are moving and elicit powerful intuitions. Of his own treatment at the hands of South African "security" forces, for example, he says that

³ Albie Sachs, "Four Tales of Terrorism," in this volume.

It was not a hypothetical situation of the kind that some academics conjure up when discussing the costs and benefits of the government using torture. And, as in 99.9 per cent recurring of cases where forms of torture are used, there was no ticking bomb nearby when I collapsed on the floor, they poured water on me, and they lifted me up. I still remember those thick, heavy fingers prying my eyes open. I collapsed again, . . . [aware of their] triumph as they were now breaking through my resistance. [page reference]

I think, however, that the emotional force of this passage derives not from its reference to an actual rather than a hypothetical example, but from the fact that it describes the torture of an innocent and indeed unusually admirable man by agents of a cruel and unjust regime acting in the service of evil ends, such as, in Sachs's words, "hegemony, dominance, power, control, mastery," and "White supremacy." [page references] This and other similar passages could be equally effective if the scenes they describe were merely hypothetical, or fictional. *Waiting for the Barbarians*, a novel by the South African writer J.M. Coetzee, has passages describing the torture of the innocent that are even more harrowing and evocative, which is hardly surprising given Coetzee's brilliance as an imaginative writer.⁴

The effectiveness of Sachs's descriptions would be lost, not if they referred to hypothetical circumstances, but if the victim were clearly *not* innocent and the aims of the torturers were clearly *just*. What would one think if the words quoted above were not those of Albie Sachs reporting what had been done to him by the guardians of Apartheid but were instead the recollections of a torturer and jailer employed by the Apartheid regime who had been beaten and deprived of sleep to force him to reveal the location of the keys necessary to release ANC members from cells in which they were being confined and tortured? Would one then feel the same identification with and sympathy for the victim, or the same outrage or indignation against the perpetrators? My own intuitive reaction would be quite different in that case. One could still

⁴ J. M. Coetzee, *Waiting for the Barbarians* (New York: Penguin, 1982).

argue that beating the jailer would be wrong even in those circumstances, but one would have to give an argument rather than simply counting on the reader's emotional reaction to deliver the desired conclusion.

Hypothetical examples, even when used as a means of understanding the most serious of moral issues, have been deployed by philosophers at least since Plato appealed to the ring of Gyges, which makes its wearer invisible, in his discussion of justice. What hypothetical examples can do that historical examples seldom can is to filter out irrelevant details that can distract or confuse our intuitions, thereby allowing us to focus on precisely those considerations that we wish to test for moral significance. Suppose that one is curious about whether a certain factor is morally significant in a certain specific way—for example, whether the intention with which a person acts can affect the permissibility of her action. It may happen that reflection on intention in the abstract proves inconclusive. One might then devise a pair of hypothetical examples in each of which an agent goes through the same series of physical movements and in which consequences of those movements are identical. The *only* difference is that in one case the consequences are intended as a means whereas in the other they are unintended but foreseen side effects. Suppose that a large majority of people from a variety of cultures judge that the agent who intends the bad consequences acts impermissibly while the agent who merely foresees them acts permissibly. That is at least *prima facie* evidence for the view that an agent's intentions can affect the permissibility of her action. Yet if one had sought to elicit people's intuitions about a pair of actual historical examples, it would have been inevitable that people would have been influenced by irrelevant historical associations, distracted by irrelevant details, or guided in their evaluations by morally relevant differences between the two cases having nothing to do with the agents' intentions. The value of hypothetical examples is that they can exclude all such features that are irrelevant to the purpose of the example.

When one understands what hypothetical examples are designed to do, one can see that the ticking bomb case is an entirely respectable philosophical tool. It is relevantly similar to thousands of other hypothetical examples that have appeared in the work of moral philosophers

in recent decades and that most philosophers regard as legitimate components of philosophical arguments. It has no features that are not characteristic of the majority of hypothetical examples in moral philosophy. It is no different in relevant respects from the familiar trolley cases, transplant cases, examples comparing and contrasting terror bombers and tactical bombers, and so on. It is, if anything, more realistic than most.

An early instance of the ticking bomb case appeared in an article published 43 years ago by Michael Walzer called “Political Action: The Problem of Dirty Hands.” Walzer describes a political leader who

is asked to authorize the torture of a captured rebel leader who knows or probably knows the location of a number of bombs hidden in apartment buildings around the city, set to go off within the next twenty-four hours. He orders the man tortured, convinced that he must do so for the sake of the people who might otherwise die in the explosions—even though he believes that torture is wrong, indeed abominable, not just sometimes, but always.⁵

Walzer uses the example not so much to defend the claim that torture can be permissible in extreme circumstances but to explore the way in which political necessity may require leaders to violate the constraints of ordinary individual morality.⁶ The main reason I mention this use of the case is to call attention to its impeccable pedigree. Walzer is notoriously averse to the use of unrealistic hypothetical examples in philosophy. The subtitle of his classic work on just war theory is “A Moral Argument with Historical Illustrations.”

In Walzer’s presentation, from which I have quoted only the essential part, the example is quite detailed and realistic. It raises the question whether what the leader did in authorizing the

⁵ Michael Walzer, “Political Action: The Problem of Dirty Hands,” *Philosophy and Public Affairs* 2 (1973): 160-80, at 167.

⁶ According to Christopher Finlay, Walzer has got it backwards, at least from a Humean point of view. Finlay argues that a moral sentiment account of morality, such as that of Adam Smith or Hume, implies that because our sympathy would be with the potential victims of the explosions rather than with the rebel, morality actually requires the torture of the rebel while political principles – those necessary for regulating the social order—oppose it. See Christopher Finlay, “Dirty Hands and the Romance of the Ticking Bomb Terrorist: a Humean Account,” *Critical Review of International Social and Political Philosophy* 14 (2011): 421-42.

torture was morally permissible. No one, so far as I am aware, objected to the case when it appeared in this well known article. Subsequent writers have added certain stipulations that make the example less realistic but more serviceable for the purpose of moral reflection: for example, the bombs have become a nuclear bomb, the rebel has been relabeled a terrorist, it is known that the terrorist planted the bomb and thus knows where it is, torture offers the only hope in the short time remaining before the bomb explodes, of inducing him to reveal its location, and so on. The example has received a lot of discussion, for the simple reason that it *does* seem to most people that, in *these* conditions, it is morally permissible for the leader to have the rebel (or terrorist) tortured.

The central deficiency of the ticking bomb case is not that it involves uncertainties that would in practice be only probabilities, but that it often tempts people to accept a mistaken explanation of why it would be permissible to torture the terrorist in these conditions. It would be more illuminating if the harm to be prevented were not cataclysmic. Suppose that instead of having planted a nuclear bomb, the captured terrorist has ordered his subordinates in an unknown location to torture a single innocent person. The aim of torturing the terrorist is to force him to disclose that location so that the further torture of this innocent victim can be prevented. We might even stipulate that, in order to be effective, the torture inflicted on the terrorist has to be more severe than that which his subordinates will otherwise inflict on their victim. It could still, I think, be permissible to torture him. If so, that would show that the justification is not a necessity or lesser evil justification—that is, it is not that the harm done to the terrorist would be significantly less than the harm that would thereby be prevented. Rather, the justification in this case is *liability-based*. It is that the terrorist's own moral responsibility for the fact that *someone* must suffer torture makes it the case that, as a matter of justice, *he* should be the one to suffer the harm that he has made unavoidable. This is a claim about justice in the *ex ante* distribution of harm, not a claim about greater and lesser harms.

It is important to stress this point, for the same considerations do the same normative work in the original ticking bomb case. The primary justification for torture in that case too is

that the terrorist has, through his own culpable action, made himself liable to be tortured in defense of his innocent potential victims. But the stipulation that there are thousands of such potential victims misleads us into thinking that it is the comparative badness of the consequences of the two possible courses of action that is doing the work of justification. Sachs himself, who suggests that such examples call for “cost/benefit analysis” [page reference] and “weighing up profits and losses in an analytical book-keeping manner,” [page reference] is among those who seem to have been misled in this way.

What the ticking bomb case shows is that, whatever we may want to say, most of us do not believe that there is an absolute moral prohibition of torture. This should not be surprising, for most of us are not absolutists about the moral prohibition of *any* general *type* of action, or act-type. Very few people, for example, are absolutists about the impermissibility of killing innocent human beings or even the *intentional* killing of innocent human beings. Most of us accept that at least one or two types of killing can be permissible in certain circumstances—for example, killing in self-defense, killing enemy combatants in a just war, capital punishment, abortion, suicide, or voluntary euthanasia. One challenge facing those who hold that torture is absolutely morally prohibited is to reconcile that view with the belief that killing can be permissible.⁷

It can, indeed, be illuminating to test proposed explanations of why torture is supposedly always morally wrong by considering whether they would be plausible when applied to killing. Consider, for example, what Sachs says would have happened to ANC members had they decided to torture their torturers: “[it] would have meant that we had become like them, that we had become gangsters and crooks and thugs—for a more noble cause to be sure, but in the end

⁷ Absolutists who believe that both torture and the intentional killing of innocent human beings are absolutely prohibited obviously have no problem of consistency here. But they face other problems, which I will briefly discuss later, that are even more serious and intractable. For various general objections to moral absolutism, see Jeff McMahan, “Torture in Principle and in Practice,” *Public Affairs Quarterly* 22 (2008): 111-28 at 111-14. On the comparison between torture and killing, see F. M. Kamm, *Ethics for Enemies: Terror, Torture, and War* (Oxford: Oxford University Press, 2011), chapter 1.

no different from them, only stronger. Our souls would be like their souls, and our inhumanity would be inseparable from their inhumanity.” [page reference] Many people have said the same about those who support or participate in one form of killing: namely, capital punishment. They claim that those who execute murderers thereby become murderers. But this is a paradigm instance of begging the question. No one other than the strictest pacifist believes that a police officer who kills a murderer when this is necessary to prevent him from killing an innocent victim thereby becomes a murderer, or morally like a murderer.

Sachs also says that he and the other members of the ANC “were making unambiguous statements about the kind of people we were, what we were fighting for, and what our morality and core values were about. . . . They did not want to belong to an organization that used torture.” [page reference] It is tempting to echo this claim by affirming that we do not want to be members of an organization that kills people. But because it can sometimes be permissible, or even morally required, to kill people, there is nothing shameful, for example, in being a soldier in an army that intentionally kills people who are liable to be killed in the course of fighting a just, defensive war. It makes a substantial moral difference whether the people we kill have made themselves morally liable to be killed. Similarly, it makes a great difference whether a person we torture is morally liable to be tortured.

Opponents of the ticking bomb example often object that this last claim presupposes that we can know when a person is liable to be tortured, whereas in practice we can never be certain of this. But if this is a serious objection to defensive torture, it must also be a serious objection to *all* forms of defensive harming. This is because there is in practice always some uncertainty about whether the target of defensive harming is liable to be harmed. It is always possible, though perhaps highly unlikely, that he poses no threat at all, or that he is not morally responsible for the threat he poses (which on my view, though not on certain other views, would exempt him from liability), or that harming him is unnecessary or would be ineffective or

disproportionate.⁸ In all these cases, I believe, the target of our attempted defensive action might not be liable to be harmed. But this unavoidable uncertainty about liability does not entail that defensive harming can never be justified. It means only that defensive harming always involves a moral risk. This risk is generally much greater in the case of torture, though I and other revisionist just war theorists believe that uncertainty about liability to be killed in war is much greater than people have hitherto supposed.⁹ In any case, this difference between torture and other forms of defensive harming is only a difference of degree, not a difference of kind.

Finally, consider Sachs's story of Sergeant Benzien, the torturer who was reduced to weeping when he was forced by the Truth and Reconciliation Commission to confront in a vivid way what he had done. [page reference to Sachs] The conclusion of that narrative would be equally comprehensible if the protagonist had not been a torturer in the service of a repressive regime but had instead been a soldier who had killed enemy combatants in the course of a just war. Suppose that such a veteran were placed in a room with the parents, wife, and children of a man he had killed in hand-to-hand combat. Suppose he were then asked to demonstrate exactly how he had killed the enemy soldier—how, for example, he had thrust his bayonet into the young man's belly and watched him writhe in agony and die. He would probably weep as well. But that would not show that what he had done was wrong. (If this is too fanciful – too hypothetical – for one's taste, ask a veteran of a just war about the people he killed. Ask him, for example, what his feelings were when he searched the body of an enemy soldier he had killed for identification and found photographs of his victim's wife and children.¹⁰)

The main reason why killing can be morally justified far more often than torture is that killing is often necessary for self- or other-defense against a threat of wrongful harm, while

⁸ I argue that moral responsibility is necessary for liability in *Killing in War* (Oxford: Clarendon Press, 2009), section 1.4. For a classic statement of the view that moral responsibility is not necessary for liability, see Judith Jarvis Thomson, "Self-Defense," *Philosophy and Public Affairs* 20 (1991): 283 – 310.

⁹ See, for example, Seth Lazar, "Risky Killing and the Ethics of War," *Ethics* 126 (2015): 91 – 117.

¹⁰ For general discussion, see Lt. Col. Dave Grossman, *On Killing: The Psychological Cost of Learning to Kill in War and Society* (Boston: Little, Brown, 1995).

torture is very seldom necessary for a defensive purpose. In part this is because in most cases a person cannot be tortured unless he has already been subdued or incapacitated and hence is incapable of posing a threat. But there are ways in which torture can be purely defensive. One way is illustrated in the ticking bomb case, in which torturing the terrorist can be a means of defending his potential victims from the effects of his previous rather than present or future action. But torture can also be defensive because pain can be a means of incapacitation.

Once it is conceded that torture, like killing, can be defensive, and that it can be less bad to be tortured than to be killed (though there are of course also forms of torture that are worse than being immediately killed), it becomes implausible to claim that a person can be liable to be killed for a purely defensive purpose but not to be tortured for a purely defensive purpose. At present, of course, it is rare that one can incapacitate a person for defensive reasons solely through the infliction of pain. This is because, again, a person must normally already be incapacitated for one to be able to cause him to suffer crippling pain in a way that does not otherwise disable him. But it is possible that a device could be developed that would act at a distance to inflict incapacitating pain on an individual without causing any bodily damage. Such a device might merely extend technologies that are already deployed in the manufacture of tasers.

Imagine that such a device exists. And suppose that an agent of a vicious, repressive regime has captured a member of the resistance, taken him to a tiny island 100 yards offshore, and begun to use a device of this sort to compel him to reveal information about his confederates. Other rebels arrive at the mainland shore but realize that it will take them an hour to get a boat to take them to the island. They are armed not only with long-range rifles but also with a device identical to the one that the government agent is using. Like the guns, the device can be used with accuracy even at a distance of 100 yards. They therefore have three options: they can allow their innocent comrade to be tortured for an hour, they can kill the torturer, or they can

incapacitate him by torturing him with the device—that is, by causing him to suffer disabling agony for an hour, until they can reach the island.¹¹

Assume that it would be permissible for them to kill the government agent if that were the only way to prevent him from continuing to torture their innocent comrade. This is a reasonable assumption; most of us accept that it would be permissible to kill a person in self-defense if that were the only way to prevent oneself from being wrongly tortured by him, and there is no reason to suppose that the same justification would not extend to third parties. Because he is culpably violating his victim's fundamental moral rights, there is no basis on which he could claim to be morally immune from defensive action. Torturing him for an hour would be proportionate, since the harm done to him would be no worse than the harm that he would be prevented from inflicting on his innocent victim. Torturing him would thus satisfy the necessity constraint while killing him would not. The necessity constraint forbids the use of a more harmful means of achieving an end if there is an alternative means that would be equally effective but less harmful. And it is uncontroversial that for a person who would otherwise have a long and happy life in prospect, it would be worse to be killed than to suffer agony for an hour, especially if the torture would not cause permanent physical damage. If killing the government agent would be permissible on grounds of liability if torturing him were not possible, then given that torturing him would be equally effective but substantially less harmful to him, it must be permissible to torture him. (Sachs himself endorses violence that is discriminate and proportionate: "The enemy was not a people, a population, but a system of injustice. When we used violence in our challenge to that system of injustice, it had always to be directed at the physical power and the structures of domination of that system, not at civilians." [page reference]) Because killing the person would be more harmful to him than torturing him, killing him would be wrong. But it would also be wrong to do nothing, allowing him to continue to torture his innocent victim. In the circumstances, therefore, torturing him is not only permissible

¹¹ This example is based on a simpler one given in "Torture in Principle and in Practice," p. 120.

but morally *required*—that is, it would be impermissible *not* to torture him. I think this is not only true but *obviously* true.

Those who believe that torture is absolutely prohibited may claim that merely *disabling* the government agent with pain does not constitute torture. But that would be idle, for no one would deny that what he is doing to his victim is torture; yet what he does to his victim is precisely what the rebels do to him. Both he and they intentionally inflict crippling pain as a means. He does so as a means of extracting information while they do so as a means of preventing him from continuing. The only difference is that the government agent inflicts pain as a means of making a person *do* something while the rebels inflict pain as a means of making a person *not* do something. He does so to manipulate a person's will while they do so to disable a person's will. It would be hard to make the case that the former is subject to an absolute prohibition while the latter is not.¹²

This example, like the ticking bomb case, is hypothetical. But both these examples help us to get clear why torture is objectionable. They help us, in particular, to see that torture cannot plausibly be thought to be absolutely morally prohibited. When its use is necessary to defend the innocent from the guilty, and would be proportionate, torture can be morally permissible. But it does not follow that it can be permissible as a matter of policy, or that it should be legally permissible. As I noted, instances in which it can be used defensively are rare. It is difficult to find actual cases in which torture has been successfully used against a known wrongdoer to prevent a grave wrong for which the wrongdoer would have been responsible. Yet the use of torture has been pervasive throughout human history. But what this means is that it has almost always been used impermissibly. It has and continues to be used almost exclusively against people who are morally innocent in support of aims that are unjust. This is why, as I have

¹² Those who are tempted to try to make that case might appeal to the analysis of the wrong involved in torture developed by David Sussman in "What's Wrong with Torture?" *Philosophy and Public Affairs* 33 (2005): 1-33.

argued elsewhere, torture must be absolutely prohibited in law.¹³ It is far more important to deter the torture of the innocent than it is to provide a legal privilege for the use of torture in those very rare cases in which it might be morally justified.

This ground for prohibiting torture in law may seem merely pragmatic, or insufficiently “principled.” Sachs refers to the absolutist position that he and others take on torture as “principled.” Yet the reasons he cites usually turn out to be matters of expediency—for example: “if in the 1980s and earlier, we had not taken principled positions on questions of . . . what methods of struggle were legitimate and how torture was to be prohibited, we would not have achieved what many today regard as the most progressive Constitution in the world.” [page reference] What, then, are the principles? Sachs makes a variety of claims.

[The delegates at the ANC conference] insisted that there be very clear standards and that absolutely no torture be used in any circumstances, whatever the euphemism used. . . . The young soldiers—and the not-so-young lawyers—were making unambiguous statements about the kind of people we were, what we were fighting for, and what our morality and core values were about. [page reference]

Whoever you are and whatever your motivation, whatever your cause, there are some things human beings just don’t do to other human beings. [page reference]

[W]hat is at stake is precisely the deep morality of the society and the role that respect for human dignity plays at its foundation, and that if respect for human dignity has a price beyond rubies, then the scales of any balancing process could never come down on the side of torture. . . . [page reference]

¹³ Jeff McMahan, “Torture, Morality, and Law,” *Case Western Reserve Journal of International Law*, 37 (2006): 241-48; and McMahan, “Torture in Principle and in Practice.”

Apart from the vague notion of human dignity, which here as in most of its other uses is merely a rhetorical gesture, Sachs's statements are devoid of substantive content. Anyone, no matter what his or her substantive moral beliefs might be, could make claims of this sort with perfect sincerity in support of those beliefs. The white South African defenders of Apartheid could echo Sachs's words by saying that "we are not the kind of people who can live on terms of equality with people lower than ourselves; rubbing shoulders with primitive, unwashed beings is not consistent with our morality. Legal and social equality between superior and inferior races is against the core values of our society." In opposing resistance to the Nazis by means of war, absolute pacifists could have said, with Sachs, that "there are some things, such as using explosives to rip their bodies apart, that human beings just don't do to other human beings." Those who oppose the legalization of voluntary euthanasia can say the same. Those who wish to impose an absolute prohibition of abortion sometimes appeal to "the deep morality of the society and the role that respect for human dignity plays at its foundation." But none of these assertions provides a *reason* for accepting the beliefs in support of which they are made.

The motivation of those philosophers who have exercised their ingenuity in producing objections to the ticking bomb case is understandable and admirable: they want to deny any support to those who defend the practice of torture as a matter of law or policy. In this they are right; it is the duty of every decent person to oppose the legalization or institutionalization of torture. But that is not because torture is absolutely morally wrong. It is instead for the reason I gave: that it is morally far more important to prevent the torture of the innocent than it is to accord a legal privilege for torture in those exceedingly rare cases in which it would be morally permissible.

Even as a tactical matter it seems wiser to base one's opposition to the practice of torture on such grounds as this than to persist in claiming that morality rules out torture absolutely. For the continued assertion of that claim is likely to do little more than to convince the defenders of torture that their opponents are obtuse or naïve. The ticking bomb case is, after all, highly effective in convincing people that torture is *not* absolutely wrong. Tactically, then, it is better

simply to grant that and to focus one's efforts instead on explaining why the fact that torture *can* be morally permissible in certain rare circumstances is largely irrelevant to matters of law or policy in a world in which vicious rulers clutch at every possible rationale for the use of torture in the service of unjust ends. It is better, if I may put it this way, to base political opposition to torture on good arguments rather than on bad ones.

2. Normative theory

Having indicated why I think that appeals to hypothetical examples such as the ticking bomb case are legitimate and can lead to conclusions that are more principled than those derived from the kinds of considerations cited by Sachs and others who reject the use of such examples, I will in this section offer a few remarks about matters of normative theory that seem relevant to understanding the morality of torture. I begin with some further observations about moral absolutism.

As I have indicated, the ticking bomb case is an effective counterexample to moral absolutism about torture. But the intuitive implausibility of absolutism can be made vivid through the use of different examples as well. Consider first a hypothetical example and then a related but actual example. Suppose there is a machine that, by operating directly on the nociceptors in the brain, causes extreme agony but without causing somatic damage. It is, we may imagine, used by various dictatorial regimes to force "confessions" from those they want to execute and that it usually produces the desired result in less than half an hour. Suppose we agree that the use of this machine to inflict pain constitutes torture. But now suppose that a man wishes to prove his strength of will and his capacity for endurance. He pays the regime's torturers to use the machine on him for an hour and instructs them not to stop even if he pleads with them to do so. This man voluntarily subjects himself to protracted torture. There is a parallel here with the actual case of the journalist and writer, Christopher Hitchens, who arranged

to have himself waterboarded by professionals whose job it is to train American soldiers in techniques for resisting torture.¹⁴

If torture is absolutely prohibited, both the imaginary person who subjected himself to torture and the people who waterboarded Hitchens acted impermissibly. They violated an absolute prohibition. It is no use pointing out that the victims in these cases *consented* to be tortured, or that the intentions of those who tortured them were benign. Absolute is absolute: there are no exceptions for consent or good intentions. This is why the Catholic Church's absolute prohibition of the intentional killing of the innocent covers suicide and voluntary euthanasia as well as ordinary instances of murder. Yet it is very difficult to believe that any of these people, actual or hypothetical, is guilty of acting in a way that cannot possibly be justified, no matter what the consequences might be of not acting in that way. We might think that the self-torturer is foolish, and that Hitchens, though well motivated, acted counterproductively, but it seems absurd to suppose that any of these parties to wholly voluntary torture is guilty of grave wrongdoing. (I suggest that Hitchens's action may have been counterproductive because it seems unlikely that the best testimony to the awfulness of torture is to reveal that one regards torture as sufficiently tolerable to be endured in order to make a political point or to enable oneself to write a fascinating article.)

As I indicated, the rejection of absolutism in morality is compatible with the acceptance of absolutism in certain areas of the law. While it is objectionable for the law to criminalize action that is morally permissible, the objection can be overridden by other considerations. Law is designed to serve certain purposes—usually morally important purposes—and the reasons for pursuing those purposes are sometimes strong enough to justify the legal prohibition of morally permissible action. There are many reasons why it may be necessary for the law to diverge from morality. Sometimes the law must permit what morality prohibits. This might be because the

¹⁴ Christopher Hitchens, "Believe Me, It's Torture," *Vanity Fair*, August 2008, reprinted in Hitchens, *Arguably* (New York: Twelve, 2011), 448-54.

law simply cannot provide sufficiently precise guidance in determining what is prohibited, so that a legal prohibition would carry too great a risk of punishing the morally innocent. Or it might be because enforcement would be excessively invasive of people's privacy. Similarly, there are various possible reasons why the law must, in some cases, prohibit what morality permits. In the case of torture, the main reason, as I have suggested, is to deny potential torturers any legal pretext or rationale for action that would almost certainly be egregiously wrong. A similar claim might be advanced in favor of the legal proscription of other types of action that may on occasion be morally permissible but are much more likely in practice to be wrong, such as the intentional killing of civilians in war, punishment by execution, and preventive war.

It is worth pausing to explain the one exception that must be granted in law to avoid making the law ridiculous. This is for torture by freely given consent, as in the case of Hitchens, or soldiers who submit themselves to mild forms of torture by their own instructors to train them to withstand the more severe forms to which they may be subjected by their enemies. This is an exception that it would be difficult to exploit. The absolute legal prohibition, then, applies only to *involuntary* torture. I will take this as given in the remainder of this essay.

The implications of the claim that torture ought to be categorically prohibited in law may seem straightforward but they are not. Further elaboration is necessary. For example, even if there are no statutory exceptions to the legal prohibition of torture, should it be possible for certain defendants to be granted a defense of necessity at trial? If the torturer in the ticking bomb case had saved the city by discovering the location of the nuclear bomb through the use of torture, ought he to be judged to have had a justification of necessity and thus be acquitted?¹⁵ I do not have a firm view about this but I am inclined, for the reasons given, to think that it is more important to deprive potential wrongful torturers of the hope of a necessity defense than it is to have that option available in the rare conditions in which torture is morally justified. Perhaps the

¹⁵ Thanks to Eric Posner for making me aware of this issue during the discussion at the conference.

best that a justified torturer ought to be able to expect at trial is to be excused, or to be treated leniently at sentencing.

Practical people, who pride themselves on living in the real world rather than in an ivory tower, may think that if we know that torture cannot be permitted in law and policy, we know all we need to know. There is nothing more to be said, and it is frivolous and repellent to conduct philosophical discussions of hypothetical cases to try to determine in which torture might *in principle* be morally justified. But this suggestion, present in much of the discussion of torture, is itself irresponsible. It is not impossible that soldiers or anti-terrorist agents could find themselves in a situation in which there seems to be a compelling moral case for the use of torture. In these situations, they may know that torture is prohibited by law and that they risk severe punishment if they engage in it, yet they may still be in need of moral guidance about whether to obey the law. For there can obviously be circumstances in which it is morally permissible, or even morally required, to break the law. Since we cannot rule out the possibility that there will in practice be cases in which torture is permissible or required, moral philosophers have a responsibility to try to provide criteria for identifying such cases and distinguishing them from others in which torture might seem to be permissible but is not. In the remainder of this essay, therefore, I will examine certain considerations that are relevant to determining when torture might be *morally* justified.

As I noted earlier, most of those who have discussed the ticking bomb case, both those who defend the practice of torture and those who oppose it, have thought that what it shows, or would show if it were successful, is that there can in principle be a *necessity or lesser evil* justification for torture. That is, they have taken it to show that although there is a uniformly strong moral constraint against torture in all cases, this constraint can be overridden if the overall consequences of not torturing a person would be vastly worse, involving many more victims, than the torture itself. When philosophers and others criticize the ticking bomb case, they typically attempt to show that the conditions for a lesser evil justification, including the *epistemic*

conditions, never arise in practice, so that the constraint is never in fact overridden. This, however, cannot be shown to be true, and is *prima facie* quite implausible.

It is essential to understand exactly what a lesser evil justification is. Although it appeals to consequences, it is not a consequentialist justification. According to consequentialism, it is permissible to harm a person whenever doing so would have better consequences than any alternative course of action in the circumstances. A consequentialist could, for instance, endorse torturing an innocent person for 100 minutes if that is the only way to prevent someone else from torturing a different innocent victim for 101 minutes, assuming there would be no other relevant differences between the alternatives. A lesser evil justification, by contrast, recognizes that there is a moral constraint against harming an innocent person that must be overridden before the harm can become permissible; therefore, it requires a substantial difference between the harm inflicted and that prevented—for example, one might claim that there is a lesser evil justification for torturing an innocent person for 100 minutes only if that is necessary to prevent another innocent person from being subjected to an even worse form of torture for a significantly longer period of time, or to prevent each of a greater number of innocent people from being tortured for 100 minutes.

I refer to the victims in all these cases as innocent people because lesser evil justifications apply only to the infliction of certain harms on people who are not morally liable those harms, who have a right not to be harmed that they have neither waived nor forfeited. To say that there is a lesser evil justification for harming them is to say that their right not to be harmed remains in force but is *overridden*. By contrast, when a person has made himself liable to be harmed—that is, when through his action he has forfeited his right not to be harmed (usually by virtue of responsibility for a wrongful harm)—a lesser evil justification for harming him may overdetermine the justification but is, in effect, superfluous. It would, indeed, be a mistake to appeal to such a justification. The harm inflicted on a person who is liable to be harmed may be fully justified even if it is not the lesser evil. It is often justifiable, for example, to inflict greater

harm on a culpable aggressor if that is necessary to prevent him from inflicting a lesser harm on an innocent victim.

The strongest form of justification for harming another person appeals to the claim that through his own action a person has forfeited his right not to be harmed, at least in certain ways, by certain persons, and for certain reasons. One form that forfeiture might take is that a person may *deserve* to be harmed. The infliction of harm on the ground that it is deserved is called punishment. I find the view, presupposed by religious people who believe in traditional conceptions of Hell, that people can deserve to be tortured deeply repellent. I will therefore set aside the idea that torture might be justified as a form of punishment. I will instead assume that the only purpose for which torture might be justified is the defense of the innocent. When harming another person is justified as a matter of defense, the justification usually appeals, again, to the claim that the target of the defensive action has forfeited his right not to be harmed. But here the form that forfeiture takes is not that the victim deserves to be harmed but that he is morally *liable* to be harmed. The main difference between desert and liability is that while the infliction of *deserved* harm may be an end in itself, the infliction of *defensive* harm must be either a means or an unavoidable side effect of averting a threat, if it is to be justified as a matter of *liability* on the part of the victim. There can be no liability-based justification for harming a person unless harming that person is either instrumental or unavoidable in producing some good effect, which in cases of defense is the prevention of wrongful harm. Whereas the infliction of deserved harm is a matter of *retributive* justice and may be justified even when further harm is wholly avoidable, the infliction of defensive harm to which a person is liable is a matter of *preventive* justice and is concerned with *who* ought to suffer harm when harm is unavoidable, so that *someone must* suffer it. Preventive justice, in other words, is concerned with the *ex ante* distribution of harm when *some* harm is unavoidable.

There is a further form of justification for harming a person that combines both a liability-based justification and a necessity or lesser-evil justification. Suppose that an event, E, that would cause great harm to many innocent people can, for some reason, be averted only if a

hundred units of suffering are either inflicted on a single person or are divided and inflicted on two or more people. Suppose that person P1 bears some degree of responsibility for the threat of E, but because the degree of his responsibility is slight, he is not liable to be caused a hundred units of suffering as a means of averting this threat. To inflict that much suffering on him would be disproportionate. Suppose, however, that *if* E could be averted by the infliction on P1 of only 80 units of suffering, he *would* be liable to that. As it turns out, E can in fact be averted by inflicting 80 units of suffering on P1 and 20 units on P2, who is not liable to be caused any degree of suffering. If preventing E were sufficiently important to justify inflicting 20 units on P2 on the ground that doing so would be the lesser evil in the circumstances, then the infliction of the remaining 80 units on P1 could be justified on grounds of liability. But if this is true, it seems that there could also be a liability justification for the infliction of 80 units on P1 and a lesser-evil justification for the infliction of the other 20 units on *P1*, who, like P2, would not be liable to suffer those 20 units. We can call this form of justification a “combined justification,” as the infliction of some of the suffering that is necessary and sufficient to avert E is justified on grounds of liability while the infliction of the rest is justified as the lesser evil.

As I indicated earlier, the ticking bomb case offers intuitive support for the claim that there can be a liability-based justification for defensive torture. The terrorist in that case has through his own wrongful action made it the case that either he must be tortured or a number of innocent people will be killed. His captors must choose between these harms: they can either allow the innocent people to become the victims of his action or they can, in effect, substitute him as the victim, using his suffering as the means of saving his intended victims. In the conditions as stipulated, torturing him is not merely permissible but morally required as a matter of preventive justice. The same applies to an accomplice who did not plant the bomb but knows where it is hidden. By refusing to reveal its location, he too is responsible through wrongdoing for the necessity of choosing between torturing him and allowing innocent people to be killed.

And as I also noted earlier, critics of the example argue that in practice the stipulated certainties—that the terrorist planted the bomb and thus knows where it is and that he will

divulge its location only under torture—never obtain. Yet some uncertainties are morally more important than others. Suppose both that anti-terrorist agents are epistemically justified in believing that a terrorist attack is imminent and that their belief is true. But they lack knowledge that would enable them to prevent it. They have captured a person who there is reason to believe is a member of the terrorist organization responsible for the threat and thus has vital information about it; yet there is uncertainty about whether he is in fact associated with this organization. This uncertainty is morally quite different from that in a case in which anti-terrorist agents have good reason to believe, but are not certain, that a terrorist attack is imminent and have captured a person they *are* certain is a high-ranking member of the terrorist organization that, if there is such a threat, is certain to know about it. This person, furthermore, is *known* to have engaged in terrorist action in the past.

Suppose that in both cases the anti-terrorist agents torture the person they have captured. In the first case, the person is not in fact a terrorist at all, while in the second there turns out to be no terrorist threat. The wrong done to the person tortured in the first case is far graver than that done to the terrorist in the second, despite the fact that in neither case was the person tortured liable to be tortured. Given the notion of liability as I have presented it, the terrorist in the second case cannot have been liable to be tortured, since torturing him could have been neither a means to nor a side effect of the prevention of a terrorist attack. Yet, although there is a sense in which he has been wronged, he may be in no position to complain. For the activities of the organization in which he was involved made it reasonable for the anti-terrorist agents to believe that there was a serious risk of a terrorist attack; thus he bears primary responsibility for the dilemma they faced—namely, the choice between torturing him, when doing so may be pointless, and not torturing him, when he had made it reasonable for them to believe that that was the only way to prevent innocent people from remaining at risk of being killed by him.

There are several ways in which the recognition that a liability-based justification for torture is possible may be important. One is that it seems, in relevant cases, to undermine the objection to torture that appeals to the helplessness or defenselessness of the victim.¹⁶ Many people believe that the objection to harming the defenseless is more than that such harming is unchivalrous. It is often claimed that there is a constraint against harming those who cannot threaten us and that this constraint is at the core of common sense beliefs about the ethics of violent conflict. It is, for example, a central element of the requirement of discrimination as it is understood in the traditional theory of the just war, which requires combatants to restrict their attacks to other combatants, and never intentionally to attack noncombatants. Some philosophers have thought that the prohibition of torture can be grounded in this same constraint. They note that for a person to be tortured, he must first be incapacitated—that is, rendered defenseless. But once he is incapacitated, he can no longer be threatening; therefore harming him can no longer be strictly defensive.

But, as the earlier example of a device that acts at a distance to cause incapacitating agony shows, torture can in principle be defensive in the purest sense—that is, it can be intended to stop an attack in progress by incapacitating the attacker with unbearable pain. And the ticking bomb case shows that action taken even against a person who is incapacitated and therefore defenseless can also be defensive, at least in the sense that it can prevent the incapacitated person's *past* action from harming its intended victims. In cases in which torturing a person is the only effective way of preventing him from wrongly harming innocent people, it may be justified on the ground that the victim has made himself liable to be tortured. In such a case, it is *good* if the victim is defenseless. He has no right of defense and it would be bad if he were able to defend himself by harming his torturers. In the ticking bomb case, for example, the terrorist can avoid being tortured simply by doing what he is in any case morally required to do, which is to reveal the location of the bomb. If instead he defends himself from torture by killing his

¹⁶ See, for example, Henry Shue, "Torture," *Philosophy and Public Affairs* 7 (1978): 124-43.

captors, thereby ensuring that his bomb will not be found and will explode, he simply compounds his offense. He will then be guilty not only of murdering and maiming the victims of the bomb but also of murdering those who tried to stop him.

Those who think that the constraint against harming the defenseless is entailed by the broader constraint against harming those who pose no threat might distinguish between torturing the government agent in my earlier example and torturing the terrorist in the ticking bomb case by conceding that the government agent is liable to necessary and proportionate harm to stop him from torturing his own captive. And they might further concede that the use of the long-range torture device is proportionate. They could concede this and still deny that there is a liability-based justification for torture in the ticking bomb case. For all these claims are compatible with the traditional just war theory's insistence that the criterion of liability to defensive attack is *actively posing* a threat to others. The government agent is engaged in actively harming another, while the terrorist in the ticking bomb case cannot actively threaten anyone, since he has been incapacitated. Yet, as I have argued at length elsewhere, the traditional just war theory's criterion of liability to defensive harm is mistaken.¹⁷ It is mistaken not only because it holds people liable for posing a threat of harm that would not be wrongful but also because it holds people liable only for *posing* a threat of harm rather than for being *morally responsible* for a threat of harm. What matters for moral liability to defensive action is moral responsibility for a threat of wrongful harm, not a person's occupation of a certain position in a causal sequence leading to wrongful harm.

There is one possibly relevant difference between torturing the government agent with the long-range device and torturing the terrorist in the ticking bomb case that I have not yet mentioned. When the rebels use their long-range device against the government agent, they are merely eliminating the threat he poses, as they would do if they were to kill him instead. By contrast, the anti-terrorist agents in the ticking bomb case torture the terrorist as a means of

¹⁷ Jeff McMahan, *Killing in War* (Oxford: Clarendon Press, 2009), Ch. 1.

compelling him to act. They need him to be present so that they can use him as a means of locating the bomb. In the terms helpfully introduced by Warren Quinn, the rebels engage in harmful *eliminative* agency while the anti-terrorist agents engage in harmful *opportunistic* agency.¹⁸ As Quinn suggests, it is plausible to suppose that harmful opportunistic agency is in general harder to justify than harmful eliminative agency, other things being equal. (One might even argue that the *unintended* harming of an innocent person is harder to justify than the *intended* infliction of equivalent harm on an innocent person via eliminative agency. The evidence for this is that while most people accept that it can be permissible to kill a person who poses a lethal threat even when the person bears no responsibility for doing so, many people reject the permissibility of killing an innocent bystander as a foreseen side effect of action necessary to preserve one's own life.) Yet even if the constraint against opportunistically using a person is in all cases stronger than the corresponding constraint against inflicting comparable harm via eliminative agency, it seems clear that people can make themselves liable to be opportunistically used in harmful ways. If that were not the case, all that culpable attackers might need in order to secure immunity from defensive harm is the passage of time. The terrorist in the ticking bomb case might be liable to torture using the long-range device to prevent him from planting the bomb, but would no longer be liable to torture once he had successfully hidden it—even if the torture would be equally effective in both cases and even if the severity of the torture required after he had hidden the bomb would be less. That seems untenable.

I will conclude by briefly considering whether certain widely recognized moral constraints on harming apply when the justification for harming is liability-based—that is, when the victim of the harm is liable to suffer it. First consider the distinction between doing harm and allowing harm to occur. Most people find this distinction important in evaluating the action of someone who is causally implicated in the death of an innocent person. In most cases it is worse,

¹⁸ Warren Quinn, "Actions, Intentions, and Consequences: The Doctrine of Double Effect," *Philosophy and Public Affairs* 18 (1989): 334-51 at 344. On the relevance of Quinn's distinction to the morality of self-defense, see McMahan, *Killing in War*, 170-72, and Jonathan Quong, "Killing in Self-Defense," *Ethics* 119 (2009): 507-37.

other things being equal, to kill an innocent person than it is to allow an innocent person to die, or be killed. It is, therefore, more difficult to provide a lesser evil justification for killing an innocent person than for allowing that same person to die, if all other relevant considerations are equal. But is this true when the person who dies has made himself liable to lose his life? Suppose, for example, that a culpable attacker is rushing toward an innocent potential victim and will kill her unless he is killed first. The victim has two options. She can either kill the attacker by shooting him or she can refrain from alerting him to the presence of a land mine in his path, thereby allowing him to be killed. If the distinction between killing and letting die is significant in this context, she has a moral reason to allow him to tread on the mine rather than shoot him. Yet it seems to me that she has no such reason.

The same may be true in the case of torture. Assume that the terrorist in the ticking bomb case is liable to experience great suffering if that is necessary to prevent the bomb he has planted from killing innocent people. Then compare the standard version of the case in which anti-terrorist agents cause him great suffering with an alternative version in which they have captured him but before they can torture him a stone drops from each of his kidneys into each ureter, causing him twice the agony that a person normally suffers from a kidney stone. Suppose this agony is as great as that caused by the agents in the standard case. The agents in the variant case have a potent new analgesic that will eliminate the terrorist's pain, but they refuse to administer it until he tells them where the bomb is located. This is a case of torture by intentionally allowing great suffering to continue. If the distinction between causing suffering and allowing suffering to occur has significance in this context, the action of the agents who allow the terrorist to suffer is easier to justify—for example, they could be justified in withholding the analgesic when the number of people they would save would be too few to justify the active torture of the terrorist. This may strike some people as plausible, but I find it hard to believe.

In both these versions of the ticking bomb case, the anti-terrorist agents intend for the terrorist to suffer great agony as a means of coercing him to reveal the location of the bomb. It seems, indeed, to be a feature of the concept of torture that the infliction of suffering must be

intended. Causing or allowing someone to suffer as an unintended side effect does not seem to count as torture. Yet there may be an exception. Frances Kamm distinguishes between acting *in order* to bring about an effect and acting *only because*, or *only on condition that*, one will bring about an effect.¹⁹ The effect is intended in the first case but not necessarily in the second. If one acts in a certain way without intending to cause great suffering to someone but *only because* one will cause him great suffering, is this a case of torture? I will leave this question open.

Most people believe that it is morally more seriously objectionable to harm an innocent person intentionally than to inflict the same harm on the same person as a foreseen but unintended side effect. If this is right, it is generally more difficult to provide a lesser evil justification for intentionally harming an innocent person than for inflicting the same harm unintentionally. But we can question whether this is true when the person harmed is morally liable to be harmed. Consider a further variant of the ticking bomb case. When the anti-terrorist agents arrive at the terrorist's hideout, they discover that he has locked himself in a fortified chamber. It will take them some time to gain entry to be able to subject him to interrogational torture. They learn, however, that the location of the bomb has been written on a piece of paper that is highly likely to be in a doubly fortified safe in a room adjacent to the one in which he has barricaded himself. To cut through the metal safe, the agents must use an instrument that unavoidably produces sounds that will shatter the terrorist's eardrums, causing both excruciating pain that will persist for days and permanent loss of hearing. Suppose that this suffering and damage would actually be worse than the torture that would almost certainly be sufficient to compel the terrorist to reveal the location of the bomb. Each means of learning the location of the bomb would take roughly the same amount of time and have a roughly equal probability of success. Assuming that the terrorist is liable to be caused great suffering if that is necessary to prevent the bomb he has planted from killing innocent people, does it make a difference morally

¹⁹ Frances Kamm, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm* (New York: Oxford University Press, 2007), Chs. 4 and 5.

whether the suffering is inflicted intentionally as a means rather than unintentionally as a foreseen side effect? Given the assumption of liability, I doubt that the intentional infliction of suffering is more seriously morally objectionable than the knowing infliction of an equal or even slightly greater degree of suffering as a side effect.

The infliction of the suffering as a side effect would not, it seems, count as torture. But what that suggests is that when someone is liable to suffer a great harm because that is unavoidable if his wrongful action is to be prevented from causing great harm to innocent people, it may be no worse to torture the person than to cause him equivalent or greater harm in a way that does not constitute torture.²⁰

²⁰ I am grateful to Scott Anderson for extremely helpful comments, both editorial and substantive, on an earlier version of this paper.