

# We don't torture: Moral resolutions, temptation, and the doctrine of double effect<sup>1</sup>

*Philosophical Lecture*  
*read 3 June 2015*

RICHARD HOLTON  
*University of Cambridge*

*Abstract:* 'We don't torture', announced both Bush and Obama. But what Bush meant as a statement of fact, Obama meant as a statement of resolve. Exploring this example, this article examines how moral resolutions work to overcome the rationalisations that temptation engenders. This in turn sheds light on the nature of our moral concepts. Resolutions are typically framed using intentional notions (torture, murder, theft). Reflection on the reasons for this provides a new perspective on what is right, and what is wrong, about the doctrine of double effect.

*Keywords:* resolution, resolve, torture, doctrine of double effect, intentional, rationalisation.

'We don't torture', said George W. Bush in November 2005.<sup>2</sup> His statement was a simple statement of fact, and it was false. The US had been torturing detainees, with the express permission of Donald Rumsfeld, the Secretary of Defense, since late 2002.<sup>3</sup>

In January 2009, on his second day as president, Barack Obama used the same words. It is possible that some torture was still then going on. But even if it was, Obama's words would not have been straightforwardly false, for he was using them in a different way, as the full quotation makes clear:

<sup>1</sup>Material from this paper was also presented to audiences at MIT, Cambridge, Sheffield, Southampton, Yale and LMU. I am very grateful for the comments received on those occasions, and to Rae Langton and Judy Thomson for comments on the written version.

<sup>2</sup>At a news conference in Panama City, 7 November 2005. For a further statement nearly two years later to much the same effect see *The New York Times* 6 October 2007: 'This government does not torture'.

<sup>3</sup>According to *ABC News*, 11 April 2008, Bush admitted that he was aware that his National Security Team had discussed 'enhanced interrogation techniques' including waterboarding. So it seems likely that he was not simply speaking falsely; he was lying.

We believe we can abide by a rule that says, We don't torture, but we can effectively obtain the intelligence we need.<sup>4</sup>

The pragmatics here are complicated. In part Obama was reporting his conviction that torture was not required to obtain the intelligence that the US thought it needed. Beyond that, though, speaking as Commander in Chief, Obama's words had the force of a command not to torture, a command that was more clearly expressed in an executive order signed on the same day.<sup>5</sup> But they also had the force of a commitment, a first-person-plural version of the kind of thing that is more normally expressed in the singular: I don't lie, I don't steal, I don't cheat, I don't torture.

My focus in this paper is on this last idea, on the idea, as I shall call it, of the *moral resolution*. Resolutions are familiar enough in other domains. Temptation tends to corrupt judgement: under its lure our view of what is best will often shift. Resolutions help to prevent this. By committing, in advance, to a certain course of action, we remove the need for deliberation under temptation, and so reduce the risk of succumbing. Much the same happens, I will argue, with moral resolutions. In the heat of the moment—a provocation, a war, a major terrorist attack—judgements may be corrupted, so that some may be tempted to pursue policies—such as torture—that in a cool hour would have been seen to be wrong. It is by forming moral resolutions we can resist such slippage.

Moral resolutions are worthy of study in themselves, and part of my interest is in looking at how they work. I want, though, to do something more. I want to use moral resolutions to shed some light on the doctrine of double effect—on the doctrine that the intention with which an action is done can have some impact on that action's permissibility. In recent years that doctrine has come under sustained criticism from Judith Jarvis Thomson and Thomas Scanlon. The thrust of their criticism stems from the initially compelling idea that in deliberating whether an action is permissible, we should look outward, to the action and its consequences, rather than inward, to the actor's intentions. Thomson asks you to imagine that you are a doctor, deciding whether it is permissible to give a dying patient the fatal dose of morphine that he requests. She says:

Suppose that, pursuant to the hospital's rules, you ask me—I am the appropriate hospital official—whether it is legally permissible for you to inject the lethal dose. I reply: 'Well, I don't know. I can't tell unless you tell me what your intention would be in injecting the drug. If you would be injecting to cause death, either as means or end, then no. But if you would be injecting only to cause relief from pain, then yes.' This is

<sup>4</sup>*The New York Times* 23 January 2009.

<sup>5</sup>The executive order required all interrogation to follow the *Army Field Manual*, a document that had forbidden torture for many years.

an absurdity. How on earth could it be thought proper for the legal permissibility of acceding to turn on what the doctor intends to bring about by acceding? Surely it should turn on the patient's condition and wishes.<sup>6</sup>

In many cases asking after an agent's intention in such circumstances would indeed be absurd, and I will endeavour to give some explanation of why. Nevertheless, I think that understanding the role of moral resolutions shows that a purely outward focus—in this case on the patient's condition and wishes—is inadequate. When an agent is in a morally tempting situation that corrupts their judgement, focusing just on the merits of the case will often lead to moral disaster. Instead, they will need to go back to the resolutions they formed earlier. This will involve a focus on their intentions twice over: first, because their resolutions are themselves intentions; and second, because, since those resolutions will themselves typically be formed in terms of intentional actions, they will need to understand their current motivations if they are to decide whether they are sticking with them or not.

If it is right that effective moral actors will often need to look in to their own intentions rather than out to the world, that rather undercuts the Thomson/Scanlon argument against the doctrine of double effect. However, it does not by itself do anything to re-establish that doctrine. Even if it is true that, in situations of temptation, actors will often need to look to their intentions, it does not follow that we assess the rightness of their actions by whether or not they do so. If actors behave well without recourse to their resolutions, that is fine; we do not think that there is something missing. And even if they do look to their resolutions, why should we think that the content of those resolutions will need to allude to intentions? So if an account of moral resolutions is to underpin a defence of something like the doctrine of double effect, more will have to be said.

The final idea I explore here is that, if they are to be effective, resolutions must indeed make reference to agents' intentional actions; so they point us towards an account in which intentional actions provide the primitive bedrock upon which our moral judgements are built. Such an account does not support a literal-minded doctrine of double effect, since, in central cases, it denies that the same action can be performed with different intentions: change the intention and one changes the intentional action. But it captures what I take to be the truth in the doctrine: namely, that intention matters.

That, in outline, is my path. I start by looking at resolutions.

<sup>6</sup>Thomson (1997).

## MORAL RESOLUTIONS

Why do we need resolutions even in simple non-moral cases? Why try to bind ourselves in advance? After all, it is surely rational to make a decision using all of the available information, including that which will only be available later. If we abstractly resolve at New Year to give up chocolate, that fails to take into account the information about just how attractive the chocolate will be each time it is actually offered, information that would be available if we were just to wait to decide each case on its merits. So wouldn't it be more rational to avoid resolutions altogether?

Most of us have enough self-knowledge to realise that there is something wrong with this line of thought. We know that the chocolate can *tempt* us: it can lead us to eat it when in prospect we didn't want to, and when in retrospect we will regret what we have done. One way to understand that is in terms of a shift in *desires*. The proximity of the chocolate, whether temporal or spatial, leads us to desire it more than we do when it is not close. But that is only part of what is going on. In addition, the presence of the chocolate influences our *beliefs*. If we start to think about it, we will convince ourselves that it will give us more pleasure than we had anticipated, or that the reasons for abstaining are not really so strong, at least not on this occasion. We can, after all, always give up chocolate tomorrow.<sup>7</sup> The rules governing our behaviour are almost never absolute—almost any rule will admit of exceptions. So, with a little ingenuity, we can almost always find the room for rationalisation and self-deception.

This means that resolutions have two complementary roles to play. They serve to enshrine our earlier desires and our earlier reasoning against the competition of the *desires* that arise in the moment of temptation. But, in addition, they serve to enshrine our earlier reasoning against the competition of the *reasoning* that we would do in the moment of temptation. That is why we should not open our resolutions up to reconsideration when the temptation is upon us. The empirical evidence bears this out. Dieters who form implementation intentions to see them through tempting moments—intentions such as 'If there are left-overs, I will freeze them immediately'—are more successful at losing weight than those who simply form the goal of losing weight. When they arrive at the tempting moment, their plan carries them through—more precisely, the presence of the left-overs serves as a cue to trigger the implementation of the action, and so the food goes into the freezer with scarcely a second thought. But if they add to the implementation intention a *because*-clause—'If there are left-overs, I will freeze them immediately *because I want to remain slim*'—all of the benefits are lost. Now the presence of the left-overs serves as cue to trigger not just an action but

<sup>7</sup>See my *Willing, Wanting, Waiting* (Holton 2009: Ch. 5) for a discussion of the empirical evidence on this.

also an awareness of the reason for that action. That in turn leads the dieters to reconsider that intention, and then to find a reason not to act on it.<sup>8</sup> Again, it is because the rules governing our behaviour are not absolute that we have the leeway to rationalise.

I suggest that much the same structure applies in the weightier cases of moral temptation, although admittedly here the empirical evidence is rather thin. Here too there is temptation to do the wrong thing, and since abstract moral principles invariably admit of plausible exceptions—that is something particularists are right about—there is room to rationalise. Resolutions are what stand in the way of such rationalisation. It is precisely because of the vulnerability of moral principles to exceptions that personal moral rules are so important.<sup>9</sup>

Some forms of moral temptation are obvious. People want things even though they know that they are not entitled to them—to take them would be stealing, or would involve cheating, or fraud or the like. In such situations normally morally upright people will sometimes rationalise modestly immoral actions, but the room to do so is limited, and so the misbehaviour is correspondingly limited. People will, for instance, take a can of Coca-Cola from a communal refrigerator—that is hardly stealing—but not the less valuable dollar bill that has been left there on a plate.<sup>10</sup>

Other cases of moral temptation are rather harder to spot, whether by actor or spectator. The temptation may not be driven by desire for personal gain. It may be driven by the desire to benefit others, or simply by the desire to be effectual, to get something done when one feels powerless. Such motivations bring us back to the dismal story of the authorisation of torture by the Bush administration.

Some think that torture is *never* morally justified. But, as anyone who has taught a first-year class on the subject knows, it is far from obvious that that is the case. It is easy to construct a story in which it looks as though torture may be permissible, or even required. A staple is the ‘ticking bomb’ case: a bomb has been planted that will kill millions of innocents; the only hope is to defuse it. The terrorist who planted it has been captured, but refuses to say where it is. Time is short.

Should the terrorist be tortured in an effort to find the bomb? Opinions differ, but many think that torture here may well be morally justified; and for those who resist such a conclusion, it is easy to increase the numbers at risk from the bomb (What if it

<sup>8</sup> Wieber *et al.* (2011).

<sup>9</sup> Particularists standardly go further and say that moral principles invariably admit of genuine, rather than merely apparent, exceptions; I am sympathetic, but remain neutral on the issue here. For the canonical statement see Dancy (1983). Clearly I am not inclined to conclude from this that moral principles have no useful role.

<sup>10</sup> Ariely (2012). He presents a vivid set of empirical studies on the role of self-deception in maintaining a morally justifiable self-image, and on the constraints on it—the extent of the moral wiggle room.

were half the world?) and the nastiness of the terrorist (What if he is gloats over the destructiveness of the bomb, which he designed to cause maximum suffering?).

Nevertheless, if such a ‘ticking bomb’ had been planted by a combatant in wartime, torture would be legally prohibited. The Third Geneva Convention allows for no exceptions. Is this because the framers of the conventions never considered ticking-bomb cases or their like? Or because, whilst the framers were aware of them, they thought it clear that torture in any such case would be morally wrong? Either is possible, but it seems unlikely that such substantial moral claims would have been built into conventions that were designed to be universally adopted. More likely the framers of the conventions thought that, although ticking-bomb cases and their ilk could arise, the chance of them actually occurring, and moreover of the people involved correctly identifying them, is small. In contrast, once exceptions that allow for torture are in place legally, the likelihood that they will be unjustifiably exploited is large. Just as dieters do better if they do not reconsider their diet at the moment they are tempted to eat, so, in time of war, combatants do better if their personal rules forbid torture. Likewise legal policies do better by blocking torture in all cases, since, if it is allowed in exceptional cases it is all too easy to come up with a story that in this case things really were exceptional. Policy on torture needs to be made in times of peace, and then adhered to in times of war.<sup>11</sup>

This is certainly the view of many who have been involved in conflict. Two former Marine Generals, Charles Krulak and Joseph Hoar, wrote in *The Washington Post* that ‘Complex situational ethics cannot be applied during the stress of combat. The rules must be firm and absolute; if torture is broached as a possibility it will become an actuality.’<sup>12</sup> A failure to see this seems to be part of what happened during the Bush ‘war on terror’; or at least, that is the most charitable gloss that can be placed on those events. Let us recall what happened. The US had fought two world wars, and then wars in Korea and Vietnam, in many cases against armies that were systematically

<sup>11</sup> It is striking that all of the four of the treaties that constitute the Geneva conventions were drafted in response to war. The first, which laid down rules to govern the treatment of the wounded and sick, was drafted in 1869 after the Crimean War and the Battle of Solferino. The second, which extended the rules to cover war at sea, was drafted in 1907 as the tenth Hague Convention, after the Russo–Japanese war. The third, which dealt with the treatment of prisoners of war, was drafted in 1929 after the First World War. And the fourth, which dealt with the treatment of civilians during war was drafted in 1949 after the Second World War, a time at which the three earlier treaties were all extensively revised. In each case the treaties tried to make good the shortcomings that the wars had revealed in earlier agreements.

<sup>12</sup> Krulak & Hoar (2007). Jeff McMahan (2008) also makes the point that the theoretical possibility of cases in which torture may be justified should not lead us to put into place institutional policies that ever allow torture (see pp. 124ff). What I am adding to this institutional claim is the idea that this reflects a general feature of resolutions, applicable even in the personal case.

using torture, without itself using torture as an officially sanctioned policy.<sup>13</sup> This changed in 2002 when, under the direction of Donald Rumsfeld, it became explicit policy to reconsider the practice of interrogation, in particular to investigate the use of so-called ‘enhanced’ techniques.<sup>14</sup> The result was the endorsement of a set of methods of interrogation—including throwing people against walls, facial slaps, stress positions, cramped confinement, sleep deprivation and, most notoriously, waterboarding—that most certainly counted as torture by international standards.<sup>15</sup>

The level of self-deception that occurred once the question of whether to torture was opened up is striking: The Office of Legal Counsel (OLC), whose role it is to give legal advice to the executive, came up with a bizarrely restrictive account of what would constitute torture. If physical it would require pain ‘of an intensity akin to that which accompanies serious physical injury such as death or organ failure’; if mental it would require suffering bringing ‘lasting psychological harm, such as seen in mental disorders like post-traumatic stress disorder’. As it stands, such a definition is scarcely coherent: death is not a form of serious physical injury; it is rather something that may result from such injury; and neither death nor organ failure need be painful. More substantially, even if something along these lines were properly formulated, it would have no precedent in US or international law, as the OLC could surely have seen. Other features of the OLC advice on the legality of torture were equally deluded: an attempt to defend it via an application of the doctrine of self-defence that was clearly legally hopeless, and an insistence, later overturned by the Supreme Court, that detainees at Guantánamo fell outside the scope of the Geneva Conventions.<sup>16</sup>

<sup>13</sup>That is not, of course, to say that US personnel never tortured in those wars; but it was never official policy. For extensive documentation of the role of the UK, France, and the US in developing and deploying ‘clean’, i.e. non-marking, torture, see Rejali (2007).

<sup>14</sup>For an account of how Rumsfeld put in place a team of interrogators and lawyers who would open these questions see Sands (2008). Rumsfeld’s task was doubtless made easier by the writings of theorists—Alan Dershowitz (2002) being the most prominent—who advocated the legalising of torture in ticking-bomb-style cases. For a more recent example, see Allhoff (2012). What is striking about this latter work is that an author whose basic motivation is consequentialist seems so utterly unconcerned about the consequences that his writing might have for those who are likely to torture in circumstances in which even he would think them unjustified.

<sup>15</sup>See Sands (2008). For a compelling first-person account of how awful waterboarding is, see Hitchens (2008).

<sup>16</sup>For discussion, see Sands (2008). Again, for the purposes of this article I give the most charitable diagnosis, the one that makes those involved least culpable. I am not sure that we should be so charitable, but many do see it this way. Thus Deputy Attorney General Joseph Margolis in his decision that John Yoo, one of the OLC team, should not face disciplinary action from the Pennsylvania Bar, wrote that ‘loyalty to his own ideology and convictions clouded his view of his obligation to his client and led him to author opinions that reflected his own extreme, albeit sincerely held, views of executive power while speaking for an institutional client’. Margolis, ‘Memorandum for the Attorney General’, 5 January 2010, available at: [judiciary.house.gov/issues/issues\\_OPRReport.html](http://judiciary.house.gov/issues/issues_OPRReport.html).

The connection between the official authorisation of torture and the general treatment of US detainees, most centrally at Abu Ghraib, remains controversial. But there is much evidence that there was influence, so that, as Krulak and Hoar put it:

As has happened with every other nation that has tried to engage in a little bit of torture—only for the toughest cases, only when nothing else works—the abuse spread like wildfire, and every captured prisoner became the key to defusing a potential ticking time bomb ... any degree of ‘flexibility’ about torture at the top drops down the chain of command like a stone—the rare exception fast becoming the rule.<sup>17</sup>

Those who advocate rules that permit torture in ticking-bomb cases and the like are keen to insist that it will be carefully regulated; but that assumes that the regulators will be immune to self-deception in the face of temptation, something that the behaviour of the OLC during the Bush presidency should make us doubt.

Why did this happen? Why did the Bush administration endorse torture during the ‘War on Terror’ after a century of conflict in which it was prohibited? There may be many factors in play. I suggest a central one was that the administration succeeded in convincing itself that things were sufficiently different that any prior resolutions—whether encapsulated in legal rules or in less formal moral thinking—simply did not apply. This was a new world, with a new kind of enemy, one to be treated in a completely different way. In retrospect, that looks absurd. But that was not how things seemed to some at the time, or, indeed, how they still seem to some now.

The approach I have been developing bears some similarities to rule utilitarianism, so let me conclude this section by stressing some differences. Rule utilitarianism insists that we sometimes have utilitarian reasons to abide by rules even if the immediate result looks to be a reduction in utility: overall, we increase pleasure if we stick to a rule of always telling the truth, even if we can find particular cases in which lying brings more pleasure. An obvious difference in the approach I am advocating is that it need not be utilitarian: resolutions can serve to protect moral thinking of any kind. A prohibition on torture can be grounded in both utilitarian and non-utilitarian considerations. But, second, and most relevant here, rule utilitarianism is typically presented as a response to the difficulties involved in calculating what action is best, and this often leads to the thought that it is a policy patronisingly enforced upon the less able by the more able. In Bernard Williams’s memorable phrase, it brings with it a whiff of Government House.<sup>18</sup> In contrast, self-deception in the face of temptation is

<sup>17</sup> Krulak & Hoar (2007). Henry Shue (1978) made a similar point about the likely spread of torture a long time ago. For discussion of the links between the torture policy authorised by Rumsfeld and the treatment of detainees at Abu Ghraib see Sands (2008) and Jeremy Waldron (2005). Waldron develops the idea of anti-torture law as an ‘archetype’, that fits well with the account presented here.

<sup>18</sup> Williams (1985: Ch. 6).

something to which we are all vulnerable. While we might hope that regulators will suffer from it less than most, they are far from immune, as the Bush White House shows. They needed the framework provided by prior resolutions; but they turned their back on them.

There is a further charge that is often levelled against utilitarianism, which is that it may well be a self-effacing policy: to be an effective utilitarian agent, one would need to believe that utilitarianism is false. Here again what I am advocating is different. To refuse to reconsider a resolution in times of temptation is not to turn against the use of reason; it is merely to turn against its use under those circumstances, exactly because those are circumstances in which it cannot be exercised well. It may be that simply trying to decide whether or not these are tempting circumstances will be enough to trigger reconsideration, in which case the disposition to stick to one's resolutions cannot itself involve the exercise of reason. But even that does not entail a rejection of reason.

My argument has been premised on the idea that ticking-bomb cases are rare, and that those that can at the time be identified as such are much rarer still. But what should happen if one really were to occur? What should an interrogator do if it was clear that very many lives were at risk, and that torture was the only likely way of saving them? I have no easy answer, just as I have no easy answer for what in general one should do if it seems that breaking one's resolutions, or breaking the law, would be best. In trivial cases, like dieting or exercise, there is always a trade-off between the resoluteness needed to avoid judgement shift, and the risk of becoming inflexibly bound to a policy that should in some circumstances be broken. The real problem in such cases is the descriptive one of explaining how we can generally succeed in not reopening a question, while at the same time remaining vigilant for those cases where we should—a phenomenon which I suspect is best explained by positing two somewhat independent systems, a conscious system of deliberation, and a second unconscious alarm system, which can trigger the first when a threshold is crossed (compare the fuel gauge on a car with the warning light that signals a nearly empty tank). In many such cases it does not matter if the alarm system goes off too readily, causing an unwarranted re-evaluation: the odd indulgence is not going to cause any harm.

In cases of potentially justified torture we cannot be so glib. If a ticking-bomb case really arose, those confronted with it would have to decide what to do, and then the judiciary and the community would have to decide how to treat them in the light of what they had done. I think that they should certainly be liable to criminal prosecution, even if we conclude that what they did was justified, although of course this should be recognised in sentencing. We can only hope that such a case will not arise, since pre-emptively building provision for it into the law will make things much worse. This is one area in which we can embrace the maxim that hard cases—hard hypothetical cases—make bad law.

So much, for now, for moral resolutions. Let us change direction for a while, and examine some recent work on the doctrine of double effect.

### THOMSON AND SCANLON ON DOUBLE EFFECT

In its classical form, the doctrine of double effect is the claim that the moral standing of an action, most centrally its moral permissibility, can turn on the intentions of the agent. More specifically, it is that claim that there are certain acts with two effects (the ‘double effect’), where acting with the intention of achieving one of them is permissible, while acting with the intention of achieving the other is not. Many discussions of the doctrine have focused on reactions to particular cases. So, for instance, some have argued that dropping a bomb with the intention of destroying a legitimate military target may be permissible, even if it is known that civilians will be killed (the war is just, the target important, there is no other way of destroying it, etc); whereas dropping the bomb in exactly the same place with the intention of killing those civilians is not. Others have disagreed. Stalemate looms.

Clearly we cannot rest with simple intuitions at this level. Judith Jarvis Thomson and Thomas Scanlon have tried to undercut the doctrine of double effect more fundamentally while explaining its apparent appeal. Their arguments are my concern here. I start with Thomson.

#### **Thomson**

Thomson makes use of a distinction between the evaluation of agents, and the determination of what is permissible: ‘the question whether it is morally permissible to do a thing is just not the same as the question whether the person who does it is thereby shown to be a good person.’<sup>19</sup> When it comes to evaluating agents—to deciding whether they are kind, or generous or principled—it is quite right to take into account their mental states, including the intentions under which they act. In contrast, when we determine the permissibility of an action we should not, as a matter of course, look to the intention of the agent who performs it. Sometimes intentions will be relevant—they may be relevant, for instance, in determining whether someone has made a promise, and that will be important in determining what is subsequently permissible. Less directly, intentions may give information about how wholeheartedly or carefully the agent will pursue the action, and that too might be relevant to its permissibility. But in the absence of special features like these, the proper evaluation of the

<sup>19</sup> Thomson (1997: 517).

permissibility of an action will proceed without reference to intentions. Proponents of the doctrine of double effect, who think that intentions will in general affect the permissibility of an action, have thus run together two very different things.

Thomson gives examples to support her argument. One case: suppose we had found out that Eisenhower was sending US troops onto the beaches of occupied France, not with the intention of liberating Europe, but with the intention of getting many of them killed. Would that have been a reason for trying to get the invasion cancelled? It might have been indirectly—a general with these intentions would hardly be a reliable military planner. But suppose we could be assured that the plan would be in no way compromised. Then, argues Thomson, we would have no reason cancel the invasion.<sup>20</sup> The permissibility of the action rests on what will happen, not on the intentions with which it is performed.

Thomson's second example we have already encountered. Consider a dying patient whose agony can only be relieved by a dose of morphine so strong that it will kill him. Does the permissibility of giving the dose depend on the intentions of the doctor? As we saw, Thomson argues that it does not. In thinking that it does, we conflate the evaluation of the agent with that of the action:

If a doctor will inject her patient intending his death as an end and, moreover, wants his death only because his death will constitute revenge, then that does matter morally. But we have to be careful about how it matters. I suggest that it has no bearing on whether it is morally permissible for her to act. Whatever her intention may be, the patient, we are supposing, desperately wants her to inject the drug.<sup>21</sup>

We should agree with Thomson that there is an important distinction between evaluating the agent and evaluating the action, and that intentions may be relevant to the former when they are not relevant to the latter. And I am also inclined to agree that it is hard to see much moral difference between intending to end the patient's pain with the knowledge that this will kill him, and intending to end his pain by killing him. However, that in itself does not show that intentions will *never* make the difference between the permissibility and impermissibility of an action. If Thomson wants to extend her argument to cover *all* cases, it has to encompass, as she accepts, the case where the doctor is motivated solely by desire for revenge. Should we accept what Thomson says there?

Suppose that, rather than being a hospital official, you are the local prosecutor, and you come across compelling evidence that the doctor's sole motivation in giving the injection was revenge. Would you be justified in instigating a murder prosecution against her? It strikes me that you would; indeed, you would be failing in your job if

<sup>20</sup> Thomson (1997: 516, n.19).

<sup>21</sup> Thomson (1997: 516).

you did not. Of course, in doing so you would almost certainly be making some assessment of the doctor's character. But I take it that that would not be the ground of the prosecution. We prosecute on the basis of the wrongness of the action, not on the defectiveness of the character. Killing from revenge is murder, and that is about as wrong an action as you can get.

Things might look somewhat different from the perspective of the hospital official, though not, I think, fundamentally so. It is a little difficult to fill out the case, since it is hard to imagine that any doctor would go to an official to ask if it is permissible to kill a patient from revenge: it is rather obvious what the answer would be. Suppose then that the official independently knows of the doctor's motivation, and is faced with the question of whether to let her go ahead. If the official's primary concern is the good of the patient, and this can only be achieved by allowing the doctor to give the injection despite her vengeful motivation, then perhaps this is the course that the official should approve. But it is hard to imagine that such advice could be given without qualm, even putting aside any concerns the official might have about becoming an accomplice. If there were any other way of achieving the good end—changing the roster of doctors, for instance, so that a compassionate one would be in charge of the patient—then surely that would be the course to take.

There would, I think, be fewer qualms in the imaginary Eisenhower case: Thomson is surely right it would have been unthinkable to call off the invasion on discovering the nature of Eisenhower's intentions. A sufficient explanation of the difference may simply be that the outcome of the invasion was so much more important than any question of Eisenhower's misdeeds. A second factor is that, whilst Eisenhower was at the centre of planning the invasion, it wasn't just his project. If it turned out that everyone involved in its organisation, from Roosevelt and Churchill down, was motivated primarily by a desire to kill Allied troops, then we might find some pressure to think rather differently; but to imagine that is to imagine a situation so bizarre that it is hard to know what to think.

This is a further difference. The vengeful doctor presents us with a clear case of murder. During a war, in contrast, it is much harder to say what is and is not murder. Of course there are murders committed by combatants during wars, on non-combatants, on combatants who have surrendered and the like. But once killing becomes morally acceptable, as it does in a legitimate war, our conception of murder becomes much harder to apply. Even a killing that is parallel to the doctor case, motivated entirely by considerations of personal revenge, is less obviously murder when the two protagonists are combatants in a legitimate war: the war provides a justification that the individual would lack. We might describe Eisenhower's motives as murderous in the envisioned circumstances, but this is more of a rhetorical move than a substantial one.

Is it important whether we can see an act as an instance of murder? I think it is—just as it was important in our earlier discussion whether an act was an instance of torture. We will return to this issue a little later. But first I want to introduce Thomas Scanlon's considerations on the topic, and to say some more about the general idea that questions of permissibility of action should look outward and not inward.

## Scanlon

In a generally very sympathetic discussion of Thomson's approach, Scanlon objects that the crucial distinction is not between the evaluation of an action and the evaluation of an agent: his worry is that when we are concerned with particular actions, we can recognise that an agent is good, and that their motivation is good, but still think that they have made a morally bad decision. So in place of Thomson's distinction, Scanlon wants to draw a new one between the role of principles as *guides to deliberation* and their role as *standards of criticism*; or, in short, between the *deliberative* stance and the *critical* stance.<sup>22</sup>

The basic idea is this: when we assess the moral standing of a possible action there are two different things that we might be doing. The first, taking the deliberative stance, involves us in looking, as if from the point of view of the actor, at the considerations for and against that action. That is what actors themselves do, but it is not limited to them; it is also what we do when we give them advice. The second, taking the critical stance, involves assessing the way that an agent goes about deciding on a particular action. Scanlon's contention is that when we take the critical stance in this latter way, intentions are relevant. But when we take the deliberative stance, they are not. When an agent deliberates what to do, they cannot look to their own intentions as relevant factors, and nor can we cite their intentions in giving them advice. (We put aside cases where the intention makes a difference to the obligations that the agent is under: for instance, where they have made a promise.)

Furthermore, according to Scanlon, the issue of permissibility concerns only the first, deliberative, question. Putting these two contentions together then takes him, via a somewhat different route, to much the same conclusion as Thomson's. For while we may take their intentions into account in *criticising* an agent, Scanlon holds that the proponent of the doctrine of double effect makes a mistake in saying that intention is relevant to the *permissibility* of the agent's action.<sup>23</sup>

<sup>22</sup> Scanlon (2008: 20–5).

<sup>23</sup> At times Scanlon seems to *identify* the deliberative question with the question of what is permissible: 'As guides to deliberation, moral principles answer a question of permissibility: "May one do X?"' (2008: 22). I am not happy with phrasing this entirely in terms of permissibility, which I think has a firmer place in a legal setting than in much ordinary moral thought. We might think, from the deliberative position,

There is something plausible in this, an insight that Scanlon's account shares with Thomson's. Surely agents cannot make actions permissible just by forming the right intention: they cannot bootstrap them into permissibility. Returning to our earlier example, the bomber pilot who wanted the civilians dead could surely not think 'It is impermissible to drop the bomb intending the deaths of civilians, but it is permissible to drop a bomb with the intention of destroying the military target, so that's what I'll intend' and thereby make the action permissible. But then equally, we might think, having formed a certain intention in the past should not make any difference. If an action is permissible this is something that is determined by how the world stands, independent of the mental life of the actor (again we put to one side special cases like having made a promise). The focus of moral deliberators should be outward, towards the world, and not inward, towards their own states of mind. We might think worse of the bomber pilot who intends the death of the civilians, but in answering the deliberative question of whether it is permissible to drop the bomb, this is beside the point.

### LOOKING OUTWARD AND MORAL RESOLUTION

Despite the plausibility of these considerations, in the light of what I said earlier about moral resolutions, I do not think that this way of putting the point can be right. It cannot be the case that in deliberating about how to act, or in advising others on how to act, we should always look out to the world, rather than in to the agent's intentions. Very often, if I am to act well, I will need to deliberate by looking to my moral resolutions. Similarly, when we ask someone for moral advice, it may be that the best they can give is to point us back to our resolutions. 'We don't torture' may have been the best thing that the Bush administration could have been told. But resolutions are just a special sort of intention. It follows that we will need to look to our intentions.

It might be objected to this that the content of the resolution is just the moral rule, and not the intention to follow that rule, so that morally resolute agents are not really looking to their resolutions, but to the rules beyond them. On this approach, in resolving not to torture, an agent comes to believe that torture is wrong; the resolution just serves to preserve that belief. That, however, is to misunderstand the role that resolutions need to play. I suggested that the agent who resolves not to torture may well think that there could be situations in which torture would be morally justified. So

---

that it would be better not to perform some selfish act, without thinking it actually impermissible to do so. Scanlon suggests that talk of the good is only apt for evaluation and not for deliberation (24–5); again, I do not see why. But I will stick mainly with talk of 'permissible', together with its dual, 'not permissible not', i.e. obligatory, and their negations.

reminding themselves of a moral rule that might admit of exceptions would not do the work that is required; on the contrary, it may well make things worse, since it would encourage further deliberation about whether the rule should apply in this case, deliberation that might well culminate in the conclusion that it should not: that was what we saw in the dieting case. What the agent needs is the resolution, which is essentially self-regarding: *I do not torture*; or, perhaps better, since it gives the individual commitment the backing of group solidarity: *We, the members of this community, do not torture*, which was the idea I attributed to Obama.

This provides an important corrective to the idea that moral deliberation should only look out to the world. In itself, though, it provides no support to the doctrine of double effect. For the advantages given by resolutions that we have considered so far are merely pragmatic: they enable us to avoid temptation. Perhaps they also confer an epistemic advantage: they enable us to keep in touch with the moral truth. But from what we have seen so far, they do not make any difference to the permissibility of an action: their existence does not make an action permissible or impermissible.<sup>24</sup>

There is, however, a further consideration. How should we understand the content of resolutions? They do not, by their nature, have to make reference to intentional actions. While I might resolve not to murder or to torture, it seems that I could equally resolve not to kill, or resolve not to cause suffering. My suggestion, though, is that to be effective resolutions will need intentional contents. Recall that, in Thomson's hospital case, things seemed to change once we raised the possibility that the doctor was in the business, not just of killing the patient, but of murdering him. Murder is an intentional notion. Let us turn once again to that.

## INTENTIONAL CONCEPTS

When Thomson and Scanlon insist that permissibility of an action is independent of the actor's intention, I take it that they are committed to thinking that the action can be identified independently of that intention. But can we identify actions independently of intentions? Many verbs, especially those that are used in moral assessment, build in intentions: 'to murder'; 'to steal'; 'to rape'; 'to defraud'; 'to mock'; 'to cheat'; 'to torture'; 'to humiliate'. The contrast is with non-intentional verbs like 'to drop' or 'to kill'. As a first approximation, we can think of the non-intentional verbs as those that we can prefix with modifiers like 'unintentionally' or 'inadvertently' without causing bafflement. We can straightforwardly say that someone unintentionally

<sup>24</sup>Scanlon sometimes talks this way; e.g. 'what *made* the action wrong was not that I have acted for a bad (selfish) reason' (2008: 24; his emphasis).

killed, but it would be rather odd to say that they unintentionally murdered. Conversely, adding the modifier ‘intentionally’ to a non-intentional verb adds content. To say that someone intentionally killed is to say more than that they killed; whereas to say that they intentionally murdered is to say no more than that they murdered. Exactly which intentions are needed for the application of any of these verbs is often a matter of debate: does murder require that one is trying to kill, or that one is reckless about whether one will? Such debates have echoed around both legal and informal usage. But there is no serious debate that there is some intentional requirement built into the very notion of murder.

Empirical psychological findings show that moral assessment and the use of intentional descriptions are intimately tied up: people tend to use ‘higher level’, intentional descriptions when a good thing is done by someone they like (‘tipping generously’), but ‘lower level’ non-intentional descriptions when the same thing is done by people they don’t like (‘adding extra money to the bill’); the tendency is reversed if it is a bad thing done.<sup>25</sup> Moving away from intentional descriptions is a way of moving away from moral attribution altogether.

If opponents of the doctrine of double effect think that intention is not relevant to permissibility, then presumably they must think that even if we naturally characterise actions using intentional concepts when we morally evaluate them, if we want to properly evaluate their permissibility we should strip them down to a corresponding thin non-intentional concept. That assumes that we can do so: it assumes that where we have an intentional action, to  $\phi$ , there will be a equivalent description, to intentionally  $\psi$ , where ‘ $\psi$ ’ is a non-intentional verb; and it is the permissibility of  $\psi$ -ing that we need to assess. But can we strip concepts down in this way?

It may be possible in any *particular* case to give a non-intentional description of a bad action in such a way that it is clear what is wrong with it. But can we give a *general* characterisation that would serve to feature in a resolution? Consider how we might try to formulate a non-intentional resolution that would do the work done by a resolution against torture. It will not do just to formulate it as ‘We will treat our detainees as well as circumstances permit’, since clearly those who are inclined to torture will argue that these circumstances do require it. So we might go to the other extreme of specificity and say: ‘We do not inflict waterboarding on detainees’, ‘We do not deprive detainees of sleep’, etc. But, whilst admirably clear, and hence well proofed against the attempt to provide exceptions, that would be much too limited; it would just be an

<sup>25</sup> Kozak *et al.* (2006). To be clear: the authors do not identify ‘higher level’ descriptions with the intentional; they rather think of them in terms of a by-construction: one does a relatively higher level action by doing an appropriate lower level one. Nevertheless, in the examples that they give, the higher level actions are typically intentional, the lower level ones typically not (‘She tipped generously by adding extra money to the bill’).

invitation to invent new tortures. Or we might try to characterise actions in terms of their effects: ‘We do not cause pain or humiliation to our detainees’, etc. But that is too broad. Mere capture will be humiliating to many, and a certain amount of pain may be the unavoidable result of keeping people secure, and of providing the medical care that they need. The natural response at this point is to say that they should not be hurt or humiliated *on purpose*; but with which purpose? People should not be hurt or humiliated, one might say, with the purpose of torturing them, which takes us back to where we started. If it is to be effective, the resolution needs to strike a balance between being too open ended, which will lead to too much reconsideration, with all of the risks of self-deluded rationalisation, and too specific, which will get the extension wrong. It will need to be, as we might say, *sufficiently predictive* of moral wrong-doing. My claim is that only the intentional notion of torture can play the needed role here.

There is nothing, apart, perhaps, from its awfulness, that is special about torture. Though I will not argue the case here, I suggest that much the same account can be given of the other core intentional notions that feature in our moral and legal thinking: to murder, to steal, to rape, to defraud and the like. These too are intentional for a reason; and we can get an insight into that reason by looking at their role in resolutions. If they are to feature in resolutions that have the right kind of force, they will need to be intentional.

Note that I am not saying that being intentional is good enough for a concept to play this role. We can create intentional notions by simple prefixing ‘intentionally’ onto non-intentional ones. From ‘to drop’ we can create ‘to intentionally drop’; from ‘to kill’ we can create ‘to intentionally kill’. That will not do the work. Return to Thomson’s case of the doctor and the dying patient. In that case, I argued, it made all the difference if we saw the doctor’s act as an act of murder. So one resolution that we might expect people to have—one that we might hope would go without remark in any decent person—is a resolution not to murder. Is that the same as a resolution not to kill? Absolutely not. The compassionate doctor who is prepared to use euthanasia sometimes has a resolution not to murder but no resolution not to kill. Her critic might, of course, argue that euthanasia is murder; but that is a substantial moral claim, and one that such a doctor will reject.

I have argued elsewhere that these intentional concepts are prime, in the sense that they cannot be factored into a mental element—the *mens rea*—and an outcome—the *actus reus*.<sup>26</sup> I will not repeat that argument, which is modelled on Williamson’s argument about knowledge.<sup>27</sup> Let me just distinguish it from a rather more familiar claim

<sup>26</sup> Holton (2015); revised version forthcoming in Gardner *et al.*

<sup>27</sup> Williamson (2000).

with which it might be confused. It is sometimes claimed that concepts like murder are morally ‘thick’ in that they combine both descriptive and evaluative content: describing something as a murder (rather than just a killing) is to express one’s moral disapproval.<sup>28</sup> About that I have my doubts. It seems to me that a sufficiently perverse god might make it morally right for me to murder, by, say, hanging the fate of the rest of the human race on whether I will do so. Indeed, if something like moral particularism is correct, it will be possible to arrive at such cases for any concept. At most there may be some *prima facie* disapproval, but even there it does not seem incoherent (rather than utterly morally perverse) for someone to argue that murder is in general a good thing. In contrast, my claim is that we cannot deny the intentionality of murder—the idea that it has a *mens rea*—although we might, of course, dispute exactly how that is to be understood.

### DOUBLE EFFECT RECONCEIVED

If this is on the right lines, then the debate around the doctrine of double effect looks rather different. As we saw, the standard move in trying to assess it is to take apparently identical actions, and then to vary the intention with which they are performed: the actor does the same thing, but with a different intent. Then a judgement is made about whether the two cases differ in their permissibility. The defender of the doctrine says that sometimes the permissibility does differ; the opponent says that it does not. My contention has been that this approach is methodologically flawed. If actions are to be characterised in intentional terms, holding the action fixed whilst varying the intention is not in general possible. To that extent, then, both supporters and critics of the doctrine have been working with a false presupposition when describing such cases.

The picture that we have in its place is of numerous different intentional concepts that we invoke when we make moral judgements. Intentions matter because these concepts matter, and they are intentional. To that extent, the defender of the doctrine of double effect has got something right. This is not to say that we have a fixed number of such intentional concepts, and that moral reasoning simply consists of applying them. The process of moral reasoning has involved refining the ideas we have (so that a killing may be a murder if it is done inadvertently but recklessly, but will not be if done in necessary self-defence; or an act of theft is not blameworthy, though it is still

<sup>28</sup> For a set of recent essays on the idea, see Kirchen (2013). Note that neither Ryle (1968, 2009), who is often taken as the originator of this approach, nor Geertz (1973), who has done so much to develop it within anthropology, thought of thick concepts as exclusively moral in this way. Plausibly, though, they did both think of them as intentional.

an act of theft, if it is done under sufficient duress). Sometimes mere refinement is not enough: the concept of sexual harassment simply was not around fifty years ago, but we have it now. Having it has enabled us to gain moral insight that was formerly obscured, and to form resolutions—how to act, but also how to respond to others' acts—in consequence.<sup>29</sup> Other concepts that once figured large in our moral thinking have now largely dropped out: for instance, the idea of honour. Just as with science so with ethics: progress requires us to get our conceptual vocabulary right.

There is thus no guarantee that every case will be covered by some relevant intentional concepts; and even if it is, there is no guarantee that our corresponding moral evaluations will be right. Much of the controversy concerning permissibility concerns exactly the cases that are not so covered: cases, for instance, around euthanasia, or the ethics of war, where much of the debate about double effect has taken place. It could be that we have yet to develop the relevant intentional vocabulary here: advances in medical technology make that not implausible in the former case at least; perhaps the same can be said of warfare, although there, while the methods have become way more efficient, the basic issues look largely unchanged. If no such development is forthcoming, we have to do the best we can using the standard combination of intuitive responses and general-purpose principles.

My surmise will be that in such cases intentions will turn out to matter little to permissibility; those are the cases to which Thomson and Scanlon devote much of their discussion, and that loom large in discussions of double effect more generally. But we will fail to see what was right about the doctrine of double effect if we fail to see that the categories we apply in more familiar terrain are shot through with intentionality.

## CONCLUDING COMMENTS

I have taken a rather winding path: from the role of resolutions, to the nature of the actions in terms of which those resolutions should be framed, to the role of intentions in moral permissibility. I am sure that that is not the only path that I could have taken to any of these conclusions. Rather than looking at intentional notions through the lens of resolutions, I could have looked at them more directly in terms of moral evaluation. There is something, I hope, to be gained, though, in looking at a topic from a slightly less familiar direction. Thomson and Scanlon have greatly improved our discussion of double effect by focusing on deliberation and advice; my contention has been that here too intention, and resolution, have a central role to play.

<sup>29</sup> For a philosophical discussion, see Fricker (2007).

## REFERENCES

- Allhoff, Fritz (2012), *Terrorism, Ticking Time-bombs and Torture* (Chicago, IL, University of Chicago Press).
- Ariely, Dan (2012), *The Honest Truth about Dishonesty* (New York, Harper).
- Dancy, Jonathan (1983), 'Ethical Particularism and Morally Relevant Properties', *Mind*, 92: 530–47.  
<https://doi.org/10.1093/mind/XCII.368.530>
- Dershowitz, Alan (2002), 'Want to Torture? Get a Warrant', *The San Francisco Chronicle*, 22 January.
- Fricker, Miranda (2007), *Epistemic Injustice: Power and the Ethics of Knowing* (Oxford, Oxford University Press).
- Gardner, J., Green, L. & Leiter, B. (eds) (forthcoming), *Oxford Studies in Philosophy of Law*, Vol 3. (Oxford, Oxford University Press).
- Geertz, Clifford, (1973), 'Thick Description: Toward an Interpretive Theory of Culture', in *The Interpretation of Cultures: Selected Essays* (New York, Basic Books), 3–30.
- Hitchens, Christopher (2008), 'Believe Me, It's Torture', *Vanity Fair*, August, <http://www.vanityfair.com/politics/features/2008/08/hitchens200808>
- Holton, Richard (2009), *Willing, Wanting, Waiting* (Oxford, Oxford University Press).
- Holton, Richard (2015), 'Crime as Prime', *Law and Ethics of Human Rights*, 9: 181–93.
- Kirchen, Simon (ed.) (2013), *Thick Concepts* (Oxford, Oxford University Press).
- Kozak, M., Marsh, A. A., & Wegner, D. M. (2006), 'What Do I Think You're Doing? Action Identification and Mind Attribution', *Journal of Personality & Social Psychology*, 90: 543–55.  
<https://doi.org/10.1037/0022-3514.90.4.543>
- Krulak, Charles & Hoar, Joseph (2007), 'It's Our Cage Too: Torture Betrays Us and Breeds New Enemies', *The Washington Post*, 17 May.
- McMahan, Jeff (2008), 'Torture in Principle and Practice', *Public Affairs Quarterly* 22: 111–28.
- Rejali, Darius (2007), *Torture and Democracy* (Princeton, NJ, Princeton University Press).
- Ryle, Gilbert (1968), 'The Thinking of Thoughts', *University Lecture* 18 (Saskatoon, University of Saskatchewan, 1968); reprinted in his *Collected Essays 1929–1968* (London, Routledge, 2009), 494–510.
- Sands, Philippe (2008), *Torture Team* (New York, Palgrave Macmillan).
- Scanlon, Thomas (2008), *Moral Dimensions* (Cambridge, MA, Harvard University Press).
- Shue, Henry (1978), 'Torture', *Philosophy and Public Affairs*, 7(2): 124–43.
- Thomson, Judith Jarvis (1997), 'Physician-assisted Suicide: Two Moral Arguments', *Ethics*, 109: 514–15.
- Waldron, Jeremy (2005), 'Torture and Positive Law: Jurisprudence for the White House', *Columbia Law Review*, 105: 1681–750.
- Wieber, Frank, Gollwitzer, Peter M., Gawrilow, Caterina & Oettingen, Gabriele (2011), 'Intending to Lose Weight: Benefits of Why Reasoning and Implementation Intentions', ms.
- Williams, Bernard (1985), *Ethics and the Limits of Philosophy* (London, Fontana Press).
- Williamson, Timothy (2000), *Knowledge and Its Limits* (Oxford, Clarendon Press).

*Note on the author:* Richard Holton is a Professor of Philosophy at the University of Cambridge, and a Fellow of Peterhouse. He works primarily on moral psychology, ethics, philosophy of law and philosophy of language. Recent publications include *Willing, Wanting, Waiting* (OUP 2009). He is currently working on a book on moral kinds.

rjh221@cam.ac.uk

To cite the article: Richard Holton (2017), 'We don't torture: Moral resolutions, temptation, and the doctrine of double effect', *Journal of the British Academy*, 5: 309–329.

DOI <https://doi.org/10.85871/jba/005.309>

This article is licensed under a  
Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.

*Journal of the British Academy* (ISSN 2052–7217) is published by  
The British Academy—the national academy for the humanities and social sciences.  
10–11 Carlton House Terrace, London, SW1Y 5AH  
[www.britishacademy.ac.uk](http://www.britishacademy.ac.uk)

