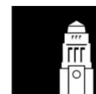
UNIVERSITY OF LEEDS

This is a repository copy of *Legitimating Torture?*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/86468/

Version: Accepted Version

Article:

Lang, G (2017) Legitimating Torture? Criminal Law and Philosophy, 11 (2). pp. 331-349. ISSN 1871-9791

https://doi.org/10.1007/s11572-015-9372-3

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Legitimating Torture?[1]

Abstract: Uwe Steinhoff (2013) defends the moral and legal permissibility of torture in a limited range of circumstances. This article criticizes Steinhoff's arguments. The analogy between ordinary defensive violence and defensive torture which Steinhoff argues for is partly spoiled by the presence, within defensive torture, of opportunistic harm, in addition to eliminative harm. Steinhoff's arguments that the mere legalization of defensive torture would not metastasize into a more full-fledged institutionalization of torture are also found wanting. As a minimal form of institutionalization, the mere legalization of torture would already be at risk of further entrenchment and growth.

Keywords: Torture; defensive violence; eliminative harming; opportunistic harming; legalization; institutionalization

Where should we hold the line against torture? We can imagine cases in which torture seems morally permissible, even required, provided we make our epistemic assumptions optimistic enough, our antagonists objectionable enough, and the numbers cataclysmic enough. Consider the notorious *Ticking Bomb* case: unless we torture a captured terrorist in order to discover the whereabouts of a huge ticking bomb we know him to have planted, thousands or even millions of innocent people will be killed. Would we really have no justification at all for torturing this terrorist?

Other imaginary cases emphasize, even less comfortably, similarities between torture and ordinary cases of defensive violence. (As we will see, *Ticking Bomb* can also be assimilated to cases with this structure.) While most of us deplore torture, we will find nothing to object to in defensive violence that satisfies the relevant tests of imminence, necessity, and proportionality. So what do such cases of defensive torture—cases in which defensive aims must be realized, out of necessity, by torture—have to teach us?

Finally, there are, of course, a few grim real-world cases which challenge the blanket moral condemnation of torture: as Uwe Steinhoff reminds us, there have been actual *Dirty Harry* cases involving child kidnapping, where torture, or the threat of torture, was deployed against the kidnapper in order to rescue a child who had been placed in a box and was in danger of suffocating.[2] These cases surely impose some pressure on those of us who would dismiss interrogational torture out of hand.

---

[1] This is a critical study of Steinhoff (2013): page references in the footnotes will abbreviate it as "ET".

[2] ET, pp. 13-14, 58, and *passim*.

Steinhoff's provocative and pugnacious book enlarges the picture sketched by his earlier articles on the subject, and argues forcefully for the moral and legal permissibility of torture in a small range of cases which, thankfully, lie largely beyond everyday experience. He does draw the line at the institutionalization of torture, which lies closer to the point of no turning back than many other thinkers would probably feel comfortable with.

Is Steinhoff's position ultimately defensible? I do not think so. He impatiently skates over certain complications in his arguments for the moral and legal permissibility of torture, and he is over-confident that the legal permissibility of torture would not be seriously abused. As I see it, the ills he attributes to institutionalization alone would also, very probably, be realized by legalization. Still, this is bold and robust work, which should be engaged with by anyone with a serious interest in these issues. Steinhoff's critical jousting sometimes inspires in him an oddly embattled, heavy-handed sarcasm—he is not someone to be relied on for lightness of touch or invariable charm—but he is an able bullshit-detector and dedicated opponent of cant. He has thrown down a gauntlet to which absolutist critics of torture must respond.

The discussion will proceed, in six sections, as follows. In Section 1, I attend to some definitional questions. In Sections 2 and 3, I trace a line of argument that takes us from *Ticking Bomb* to cases of defensive torture, and examines the dialectical role of defensive torture in Steinhoff's overall argument. Defensive torture cases are then revisited, in a more critical spirit, in Section 4. Section 5 compares and contrasts different grades of permissibility—moral, legal, and institutional—and criticizes the stability of the position Steinhoff advances. Section 6 offers a few further remarks on dehumanization, brutality, and Jeremy Waldron's notion of a legal archetype.

1. Introduction: Defining Torture

What is torture? Two features associated with torture may quickly leap to mind: the pain or suffering it involves, and the manner of its infliction. In typical acts of torture, pain is deliberately inflicted by a torturer upon a victim, who is powerless to resist it, for a certain purpose. These purposes can be highly varied: there can be recreational or sadistic torture, penal or punitive torture, and interrogational torture. (This is not an exhaustive list.) Interrogational torture is the type of torture most often discussed in the philosophical literature, for the obvious reason that other purposes seem wholly incapable of collecting any moral justification.

Steinhoff's definition of torture is simpler: "Torture is the knowing infliction of continuous or repeated extreme physical suffering for other than medical purposes".[3] He is swift to point out that this definition of torture does not require the usual accompaniments of defenceless and vulnerability. Imagine a case in which a jeweller counter-attacks an armed robber by activating a pain-inflicting device that causes anyone within a 10 metre radius (except the person wielding the device) to experience extreme pain.[4] By Steinhoff's definition, the jeweller's activation of the pain-inflicting device will count as torture, even if the robber is unrestrained and armed, and can therefore fight back. It follows, then, that you do not have to be defenceless, powerless, or restrained, in order for it to be true that you have been tortured. Here is an alternative and less fanciful example which also conforms to Steinhoff's definition:

> [C]onsider… a woman strangled from behind in her car by the serial killer who has waited for her on the backseat. The attacked woman cannot reach the gun she put into the glove compartment shortly before the killer's attack, but she has reached the cigarette lighter that has just become hot. The serial killer strangling her is determined not to let go of her because he fears precisely that she might then reach the gun. She presses the hot cigarette lighter hard onto the back of his hand; the man tries to bear the pain, tries to bear it—but finally can't and lets go.[5]

Perhaps activation of the pain-inflicting device and deployment of the hot cigarette lighter do indeed count as acts of torture. My own semantic intuitions, for what they are worth, do not clearly condemn Steinhoff's definition. The main point not to lose sight of at this early stage is that, even if the defencelessness and vulnerability of the torture victim are not necessary for the act to qualify *as* torture, these features may nonetheless help us to explain *why* torture is wrong, and perhaps *always* wrong, when it is inflicted upon individuals who meet those further typical conditions. When we add the defencelessness or vulnerability of the recipient of torture to the fact that intense pain has been inflicted upon him, we may end up with a practice whose wrongness is not to be attributed to the degree of suffering alone. This point, to which Steinhoff proves strenuously resistant, will be revisited later.

---

[3] ET, p. 7.
[4] ET, pp. 7-8.
[5] ET, p. 114. The case is presented in the context of Steinhoff's capable discussion of the "breaking of will" objection to torture emphasized by some absolutists.

2. Ticking Bombs and the Balance of Numbers

Many people are categorically opposed to torture. They are *absolutists* about torture: for them, torture is never justified, regardless of the circumstances. Absolutists think that torture is morally impermissible in all possible worlds.

Can this absolute prohibition on torture really be maintained? Let us revisit *Ticking Bomb*:

> *Ticking Bomb*
> A terrorist, *R*, has been captured with culpable responsibility for a huge ticking bomb, shortly due to explode in a large city with foreseeable casualties of up to one million innocent people. *R* refuses to disclose the whereabouts of the bomb.

Can *R* be tortured to get him to disclose this information? If the stakes are high enough (as they seem to be) and our epistemic assumptions are optimistic enough (adjust as necessary), the answer might be: yes. Some absolutists may demur on the grounds that the moral badness of torture is somehow incomparable with the moral badness of ordinary death. If so, we might appeal to the following case:

> *Torture Codes*
> A terrorist, *S*, has been captured with responsibility for the activation of a large fleet of *torture robots*, shortly due to go to work on a large city with a foreseeable torture victim tally of up to one million innocent people. *S* refuses to disclose the codes for de-activating the programme.

Torture Codes is a fanciful concoction, to be sure, but what we are looking for is a case which challenges the absolute moral prohibition on torture. To challenge absolutism, we are licensed to roam over all possible words, not just the worlds containing ordinary possibilities. If the depth of one's moral objection to torture is supposed to generate, or be reflected in, an absolute condemnation of it in all possible worlds, then *Ticking Bomb* and *Torture Codes* are cases which challenge the depth of that objection. We can also, more grimly but less fancifully, appeal to real-life cases to illustrate the torture-for-torture constraint. Take real-life *Dirty Harry* cases again, where children are kidnapped and then placed in boxes to suffocate. Steinhoff's view is that such a kidnapper is guilty of torturing the child,

as a result of gradually suffocating him,[6] so the infliction of interrogational torture on the kidnapper would at least satisfy the torture-for-torture constraint.[7]

Of course, some hard-core absolutists may resist the force of such cases. Their position will come under renewed pressure in the next section, when we see how Steinhoff enrols self-defensive considerations into the argument. Pure pacifists are thin on the ground, and absolutist opponents of torture will usually be reluctant to join their ranks. If there are cases of torture which demonstrably possess a defensive structure, then absolutist opponents of torture may either have to rethink their theories of self-defence, or consider making some concessions to the non-absolutist.

Ticking Bomb and *Torture Codes* are cases whose force can be acknowledged both by consequentialism and threshold deontology. For consequentialists, it is obvious enough that the numbers will eventually make the difference, and come to outweigh the combined badness of torture and the badness of the wider social effects of the practice of torture. The same goes for threshold deontology, to whose ranks Steinhoff assigns himself. Though deontic constraints will hold at sub-catastrophic levels of harm, these constraints will weaken when the numbers reach a catastrophic size.[8] But two further aspects of this position offer some succour to the threshold deontologist.

First, threshold deontologists are, in a sense, *protected* by the numbers, and indeed by the fanciful nature of *Ticking Bomb* or *Torture Codes*: since such cases very rarely occur in ordinary life, torture will be justified only in exceptionally rare circumstances.[9] The moral license generated by *Ticking Bomb* can therefore be comfortably partnered with a sincere and thorough-going opposition to torture as it is practised in the actual world.

Second, even in cases where it would not be wrong to use torture, the victims of torture are not properly regarded as being *liable* to be tortured: torturing them is not wrong, but they are still *wronged* by being tortured. This is because their rights not to be tortured are not erased, but simply overridden. These cases

---

[6] This type of torture will of course be sadistic, or recreational.
[7] ET, pp. 81, 85-86, 97, and *passim*.
[8] Steinhoff's "Argument from Necessity" (ET, pp. 39-45) seems little more than a recital, without much supporting explanation, of this fundamental commitment of threshold deontology. As I see it, the sense of necessity invoked in the Argument from Necessity is not to be confused with the necessity condition on defensive violence.
[9] As we shall see in Section 5, Steinhoff himself leans on the unusual nature of these cases in his argument for legalization.

involve rights infringements. It can be morally right to infringe rights, but the victims of such infringement retain their moral status insofar as they are wronged as a result of these infringements.

3. Defensive Torture and Liability-Based Cases

We can imagine still more drastic versions of *Ticking Bomb* where, faced with *R*'s recalcitrance, it would be necessary to torture *R's innocent sibling* in order to get *R* to release the information. (Though stoical in the face of his own suffering, assume that *R* is particularly close to his sibling, so that the prospect of her torture can be expected to get results.) Call this the *Extreme Ticking Bomb* case.[10] If the prospect of *Extreme Ticking Bomb* is substantially less appealing than the original version of *Ticking Bomb*, something else might be affecting our judgments in that original case, which should prompt us to take another look at it. And the something else which sets *Ticking Bomb* apart from *Extreme Ticking Bomb* may be a suspected contrast between *R*'s sibling's non-liability and *R*'s liability. *R*'s sibling is not liable to be tortured or harmed, because, as an assumed innocent, she has done nothing to make herself liable. *R*, by contrast, can be argued to have rendered himself liable to be harmed on defensive grounds by planting the bomb. So there might yet be a liability-based angle from which to approach *Ticking Bomb*.[11]

So can this suspicion be upheld? Can there be cases in which torture victims are not wronged by being tortured, i.e., cases in which they may be actually liable to be tortured? These cases would represent a more profound divergence from absolutism. And this is where cases of defensive torture make their entrance into the dialectic.

I will now reconstruct the main trajectory of Steinhoff's argument. I will organize it in my own way, and add some of my own terminology to bring out the underlying points and transitions of thought.

Compare and contrast the following cases:[12]

Poisoning 1:

X kills *Y* in order to stop *Y* from poisoning *X*.

---

[10] Even more extreme variants of *Extreme Ticking Bomb* cases can be fashioned: the selected torture victim, or victims, might include children, for example. Steinhoff's Argument from Necessity suggests that he would be prepared to countenance the use of torture in such cases.

[11] Cf. McMahan (2008), p. 118.

[12] I borrow these easy-to-present cases from Griffin (2010), p. 9.

Poisoning 2:

X tortures *Y* in order to get *Y* to give *X* the antidote to the deadly poison which *Y* has just administered to *X*.

Poisoning 1 seems to be just an ordinary self-defensive case. Self-defensive violence is permissible just in case the proportionality, necessity, and imminence conditions are satisfied. *Poisoning 2* involves torture, but what is the deep moral difference between *Poisoning 1* and *Poisoning 2* supposed to be? The structure of the two cases seems analogous. True, the timing of the attack is different, because in *Poisoning 2 Y* has already administered the poison to *X*, whereas in *Poisoning 1 Y* has not yet managed to do so. As a result, the content or mechanism of defensive response is also different: *X* kills *Y* in *Poisoning 1*, whereas *X* tortures *Y* in *Poisoning 2*. But a fundamental moral similarity in *X*'s responses between the two cases is ensured by the combination of the following features: *X*'s innocence; *X*'s aim of protecting herself from *Y*'s attack on her; and the fact that killing *Y* in *Poisoning 1* and torturing *Y* in *Poisoning 2* are necessary to the achievement of *X*'s aim. In *Poisoning 1*, *X* ensures that *Y*'s attack on her is unsuccessful by killing *Y*. In *Poisoning 2*, *X* ensures that *Y*'s attack on her is unsuccessful by extracting, through torture, the antidote to the poison which *Y* has just administered to her. These facts suggest that, when all is said and done, the two cases are normatively analogous. In both cases, the violence deployed against *Y* is justified by the combined operation of the necessity condition and imminence condition on self-defence. Call this the *Analogy Claim*.[13]

Now consider two structurally similar extensions of *Poisoning 1* and *Poisoning 2*. *Poisoning 3* appears to be a legitimate extension of *Poisoning 1*, involving other-defence, rather than self-defence:

Poisoning 3:

Z kills *Y* in order to stop *Y* from poisoning *X*.

If *X*'s actions in *Poisoning 1* are permissible, then *Z*'s actions in *Poisoning 3* should also be permissible. Now for the second extension: *Poisoning 4* appears to be a

[13] This is my term, not Steinhoff's. The same holds for the "Relative Harm Claim" and "Permissibility Claim", introduced below.

legitimate extension of *Poisoning 2*, involving other-defensive torture, rather than self-defensive torture:

> *Poisoning 4:*
> *Z* tortures *Y* in order to get *Y* to give *Z* the antidote to the deadly poison which *Y* has just administered to *X*.

If *X*'s actions in *Poisoning 2* are permissible, then *Z*'s actions in *Poisoning 4* should also be permissible.[14] But if we have managed to trace a path all the way from the uncontroversial *Poisoning 1* to *Poisoning 4*, then we have, in effect, established a striking structural similarity between ordinary cases of self-defence and cases of interrogational torture.

We have travelled a long way in a short time. Was there really enough normative fuel in our tank to get as far as this? If not, at which earlier point should the journey have been terminated? Such concerns might prompt the following objection, which I will call the *Content Objection*:

> *Content Objection*
> Hold on. *Poisoning 2* and *Poisoning 4* involve torture, not ordinary defensive violence. Isn't *that* the difference? It does not immediately follow, just because *Y* is liable to defensive violence, that he is also liable to be tortured, even if torturing him achieves the same goals as ordinary defensive violence, and even if *Y* seems just as morally objectionable in *Poisoning 4* as in *Poisoning 1*. That is to confuse the *structure* of the case, with its similar-looking parameters of imminence and necessity, with the specific type of act torture *is*, or the specific *content* of torture: the deliberate infliction of intense pain on individuals who, in interrogational contexts, are powerless to resist it. Torture, unlike ordinary defensive violence, can be properly described as cruel, sadistic and inhumane, since it is typically inflicted upon those who are powerless to resist it. Or, if we want to go in a more Kantian direction, we can emphasize the particular way in which it *uses* those on whom it is inflicted. *However* we choose to round out our moral story about the

[14] Steinhoff treats the extension from self-defence to other-defence as unproblematic. Interestingly, he also notes that, in German penal law, the generic term is "*Notwehr*", which is better translated as "emergency defence", applicable to both self-defence and other-defence (ET, p. 11). I note one or two dangers with the extension from self-defence to other-defence in Lang (2014), pp. 45-47.

badness of these properties of interrogational torture—there is certainly more to say about them—we have no business, at this early stage, in *overlooking* these properties if we are trying to gauge the moral permissibility of torture. The Analogy Claim, which pretends these further questions don't even arise, surely begs the question!

How forceful is the Content Objection? Steinhoff would resist it. His chief resource for being untroubled by it lies in the thought that being killed is often *worse* than—more harmful than—being tortured. In a particularly forthright articulation of this idea, he writes:

> Clearly, the absolutist torture opponent who, full well knowing that the victim will only be waterboarded for half an hour, kills him anyway to spare him a fate allegedly "worse than death" is quite simply morally insane.[15]

Many other passages make exactly the same point (though sometimes in softer tones). I shall call the claim that being killed is often more harmful than being tortured the *Relative Harm Claim*.

Armed with the Relative Harm Claim, let us reconsider *Poisoning 1* and *Poisoning 2*. Is it more harmful for *Y* to be killed in *Poisoning 1*, or tortured in *Poisoning 2*? Surely *Y* could intelligibly choose to be tortured, rather than killed. Even if *Y* is physically damaged and psychologically scarred as a result of torture, he may still have a life worth living, which contrasts favourably with having no life at all. On this view, it would be *better* for *Y* to be tortured rather than killed. But, if that is so, then it can hardly be *morally* worse to torture *Y* than to kill *Y*. How can ordinary non-lethal defensive torture be worse than lethal ordinary defensive violence?[16] We can also imagine the same reasoning appealing to a third party; we do not have to appeal to the preferences of the torture victim. Steinhoff imagines a case in which a dictator gives prisoner *A* the following choice: either choose one out of 10 prisoners to torture for two hours, or choose one of them to kill. *A* is forbidden to consult the prisoners' preferences. If he attempts to do so, the dictator will kill them all. On Steinhoff's view, *A* should clearly torture one of the prisoners for two hours, rather than killing him.[17]

---

[15] ET, p. 33.
[16] The trajectory of argument plotted by Hare (2014), esp. pp. 387-388, seems rather similar.
[17] ET, pp. 106-107.

Steinhoff's argument for the moral permissibly of defensive torture is constructed, then, out of reflection on the three conditions on legitimate self-defence: the imminence, necessity and proportionality conditions. As he says, in an executive summary of his main line of argument:

> My moral argument is that if injuring or killing a person in self-defense can be morally justified as long as the defense meets the so-called necessity, imminence and proportionality or no-gross disproportionality requirements, the same is true for torture.[18]

And, even more succinctly:

> Killing is worse than some forms of torture. It is sometimes permissible to kill. Therefore, it is sometimes permissible to torture.[19]

In my reconstruction of his main trajectory of argument, the analogy between *Poisoning 1* and *Poisoning 2* is secured by reflection on the imminence and necessity conditions, and yields the Analogy Claim; while the Relative Harm Claim, which is secured by reflection on the proportionality condition, manages to silences the Content Objection. As soon as Steinhoff's arguments for these conditions have been made, the argument is basically wrapped up.[20] The rest of the book is devoted to various replies and rebuttals, together with an exploration of legalization and institutionalization.

I will make four further recapitulative remarks before reviewing this territory in a more critical way.

First, with *Poisoning 3* and *Poisoning 4* on the moral books, we now have a perfectly secure *non-consequentialist* justification for torture in certain cases. We no longer need the consequentialist inflections of *Ticking Bomb*. Torture becomes a moral option even when we are out of range of the catastrophic numbers involved in *Ticking Bomb*.

---

[18] ET, p. 17. The distinction between the proportionality condition and the no-gross disproportionality condition is discussed at ET, p. 13.
[19] ET, p. 29.
[20] The main argument is essentially over and done with as early as ET, p. 19. At ET, pp. 35-38, Steinhoff also offers the "Argument from the Culpability for Creating a Forced-Choice Situation". (See Montague (1981) and (1989) for the original source of the argument.) On my view, it is better to classify the forced-choice justification, not as an entirely independent argument, but as a particular way—one among many—of grounding the defensive permissions which lie behind the self-defence justification. In any case, I won't pay any further attention to it here.

Second, and relatedly, interrogational torture has a built-in other-defensive structure. *Poisoning 4* is a form of interrogational torture. But *Poisoning 4* represents nothing more than a conceptually and morally intelligible extension of the materials that were already in place in *Poisoning 1.*

Third, Steinhoff thinks that the moral permissibility of defensive torture is basically *settled by* the combination of the Analogy Claim and the Relative Harm Claim. Nothing further, in such cases, has to be investigated in respect of the moral propriety of torture. The fact that, in typical interrogational contexts, torture involves the infliction of severe pain on an individual who is defenceless and powerless at the point at which torture is inflicted on him is not a consideration which can disturb the moral preferability of non-lethal torture over lethal defensive harm. Provided that interrogational torture is less harmful than lethal defensive harm, and provided that lethal defensive harm would not be morally ruled out by the combined imminence, necessity and proportionality conditions in the case in question, it will straightforwardly follow that torture would be morally preferable to death, which is all that is needed to secure the fact that defensive torture is morally permissible.

The fourth point is concerned with the more fine-tuned structure of this argument. Though it is the Analogy Claim and the Relative Harm Claim which seem to be doing the main explanatory work for him, Steinhoff's argument is actually dependent on a further implicit assumption, which works in partnership with these other claims. We may call it the *Permissibility Claim*: if φ-ing is permissible, and φ-ing is more harmful than β-ing, and there are no further significant moral disanalogies between φ-ing and β-ing, then β-ing is also permissible. As I see it, then, Steinhoff's conclusion is delivered by these three claims acting in combination: the Analogy Claim, the Relative Harm Claim, and the Permissibility Claim.

4. Deeper into Defensive Torture

I will consider three challenges to Steinhoff's position. Two of these are discussed directly by Steinhoff, while the third challenge emerges out of the second of them.[21]

The *First Challenge*, as I will call it, returns us to the defencelessness of the torture victim in interrogational torture. (This property of interrogational torture

[21] In the general review of recent literature offered in Chapter 6, Steinhoff discusses a large number of challenges. I make no attempt at completeness here, but will simply pick out what I regard as the most significant challenges.

was mentioned in the Content Objection.) Though Steinhoff's definition of torture does not insist on the defencelessness of the torture victim, cases of interrogational torture nonetheless exemplify an extreme asymmetry of power: the torturer is all-powerful, and the torture victim is powerless. This feature contrasts sharply with ordinary cases of defensive violence, where the attacker's normative vulnerability to self-defensive attack is explained by his active dangerousness, and is accompanied by his option of withdrawing from the attack. The property of defencelessness lies at the heart of Henry Shue's well-known moral analysis of torture's wrong-making characteristics.[22]

To discuss the First Challenge, consider *Poisoning 4*. Steinhoff's various remarks suggest the following replies to it: *Y* is still dangerous, and he is far from defenceless.[23] *Y* is actively dangerous, even if he is tied up, because he has initiated an attempt on *X*'s life which is still ongoing, and whose success has yet to be determined. *Y* is also in possession of a defence: to stop *Z* from torturing him, he can simply disclose to *Z* the required information about the antidote. Turning to other aspects of the First Challenge, it may be true of *Poisoning 4* that the torturer, *Z*, is not himself in danger, unlike the ordinary self-defensive case, but this point does not count against cases of other-defence. Besides, there are conceivable cases, such as the self-defensive *Poisoning 2*, in which the torturer, *X*, *is* in danger. There is also one further type of case in which ordinary self-defensive violence is often deemed to be justifiable, though it is directed against someone who is defenceless: these are cases involving the "innocent threat", standardly illustrated by an agent falling through the air towards the victim. The innocent threat is dangerous, but not in virtue of his agency. (This innocent threat is *falling*; there is actually nothing he is *doing*.)[24] These are forceful replies. Even if the First Challenge is on to something, it badly stands in need of supplementation.

The *Second Challenge*, as I will call it, is due to Whitley Kaufman. Kaufman's chief complaint is that, in defensive torture, the harm is not genuinely *defensive*, but *instrumental*. Kaufman sets out his central argument in the following passage:

> The doctrine of self-defense permits only force of a certain kind: force that is defensive, that wards off or literally fends off the threatened harm. It has

[22] See Shue (1978).
[23] ET, pp. 81, 83, 92-96.
[24] ET, p. 95. See Thomson (1991) for an influential discussion. Thomson's view that it is permissible for the victim to kill the innocent threat in self-defence is controversial, however: see, for example, Otsuka (1994) for an opposing view.

> never been interpreted to permit using a person—even a Culpable Aggressor—as a means to escape harm—that is, to use instrumental force… [E]ven if he were considered a "continuing aggressor," it would not follow that any force against him is permissible. There would have to be evidence that the force being used against him is genuinely defensive, not instrumental force used as a means to a further goal… The use of torture to achieve a further goal… is a classic example of instrumental rather than defensive force. It is force used a means to prevent future harm, just like using an innocent bystander as a human shield… That torture is not defensive but instrumental force can also be seen in that it is in practice irrelevant who one tortures so long as one achieves one's goal; for example, one could try to get the terrorist to reveal the information by torturing his wife or child in front of him.[25]

Applying Kaufman's reasoning to *Poisoning 4*, the central point is that the harm *Z* inflicts on *Y* is instrumental rather than defensive because it is inflicted on *Y* in order to bring about some greater good (namely, the obtaining of the antidote). Thus *Y* is being treated as a *bystander*, albeit (in this particular case) a *culpable bystander*, rather than as an *attacker*. So the argument in favour of defensive torture over-generates implications: it shows that bystanders, not just attackers, can be tortured in pursuit of a good end. These implications will be entirely unacceptable to non-consequentialists, and they point to severe limitations in the force of the Analogy Claim.

Steinhoff's reply to the Second Challenge comes in two parts, and draws on some of the materials already advanced against the First Challenge.[26] Again, I will apply the substance of his reply to *Poisoning 4*. First, *Y* is not merely a bystander, even a culpable bystander, because, by assumption, *Y* is the very agent who is posing the threat. Second, *Y* is not defenceless, for reasons that have already been rehearsed: it lies within *Y*'s power to get the torturer to desist from torturing him simply by revealing to his torturer the desired information. Thus *Y* is not, contrary to what Kaufman alleges, being punished for an earlier crime, or being treated as a culpable bystander. Since it is basically within *Y*'s control whether the torture ceases, *Y*'s position in *Poisoning 4* is relevantly similar to his situation in *Poisoning 1*, where he is just an ordinary culpable attacker: he can escape the infliction of violence on him by abandoning the attack which, it goes without saying, he was morally forbidden from making in the first place.

As far as it goes, Steinhoff's response to Kaufman's concerns is not implausible. But there is a residual challenge contained in Kaufman's critique

---

[25] Kaufman (2008), pp. 109-110. The passage is quoted in ET, pp. 97-98.
[26] ET, pp. 98-99.

which can be brought out more forcefully if it is decoupled from the actual execution of his argument. In my view, Kaufman actually diminishes the potential force of his own argument by inferring that, if the harm inflicted on the terrorist is genuinely instrumental, it can also be inflicted on someone who is not liable to be harmed. It would not immediately follow from the fact—if it were a fact—that the violence inflicted on the terrorist has an instrumental character that "it is in practice irrelevant who one tortures so long as one achieves one's goal". Perhaps the infliction of instrumental violence is still properly regulated by facts about liability. Kaufman does not justify his inference, and we will shortly see why he may have thought that it requires no justification. In any case, Kaufman does not *need* the inference. The possibility that defensive torture may be "defensive" *and* "instrumental" is enough to raise significant questions about the success of Steinhoff's argument, as we are about to see.

This takes us, accordingly, to the *Third Challenge*, which is a more refined version of the Second Challenge. It is not discussed directly by Steinhoff (though it is easy to imagine how he might attempt to tackle it). Since the Third Challenge is based on the distinction between two types of harming—eliminative harming and opportunistic harming—I need, first, to explain this distinction. The eliminative/opportunistic distinction was originally introduced by Warren Quinn in his exploration of the doctrine of double effect. Quinn makes two sets of distinctions. The first distinction is between "direct" and "indirect" types of harming. In agency where the harm is direct, "the harm comes to some victims, at least in part, from the agent's deliberately involving them in something in order to further his purposes precisely by way of their being so involved (agency in which they figure as *intentional objects*)".[27] The harming that results from indirect agency, by contrast, is "agency in which either nothing is in that way intended for the victims or what is so intended does not contribute to their harm".[28] Both ordinary self-defensive violence and defensive torture count as forms of direct harming, according to this distinction. In the deliberations of the self-defender or other-defender, the recipients of such violence unquestionably emerge as "intentional objects". Quinn's second distinction, between eliminative and opportunistic harming, is a distinction *within* the category of direct harming. This second distinction arises with respect to whether the harming agent "sees the victim as an

[27] Quinn (1989), p. 343.
[28] Quinn (1989), p. 343.

advantage or as a difficulty".[29] Eliminative harming, which regards the victim as a difficulty, "aims to remove an obstacle or difficulty that the victim presents", whereas opportunistic harming, which regards the victim, or the presence of the victim, as an advantage, "benefits from the presence of the victim".[30] Other things equal, to be harmed in an opportunistic way is worse, because it involves being used, and being used is a distinct wrong-making property.

With this background in place, we are now in a position to state the Third Challenge in full: defensive torture, unlike ordinary defensive violence, involves opportunistic harm, since it uses the presence of the torture victim to promote some other purpose, and it thus bears a wrong-making property which is lacking from ordinary defensive violence. Ordinary defensive violence is squarely eliminative, not opportunistic.

How would Steinhoff tackle the Third Challenge? Two strategies seem available to him. First, he might advance the argument that the harming is eliminative in *Poisoning 4* because *Y*'s profile in that case is still the profile of an *attacker*: in both *Poisoning 4* and *Poisoning 1*, *Y* is attacking *X*. Second, even if it were true that defensive torture involves being used, it would still be sensible for one to select it over death. Death would still look like a worse fate than certain forms of torture. But if death is both morally permissible, as it will be in certain defensive cases, and if death is worse than torture, then non-lethal defensive torture must be permissible as an alternative to lethal defensive force. So, if it would be permissible for *X* to kill *Y* in the circumstances of *Poisoning 1*, then it must also be permissible for *X* to torture *Y* in the circumstances of *Poisoning 2*. And, when we provide for the extension from self-defence to other-defence, it must also be permissible for *Z* to torture *Y* in the circumstances of *Poisoning 4*, if it is permissible for *Z* to kill *Y* in the circumstances of *Poisoning 3*.

The first of these suggested counter-arguments is dependent upon an implicit assumption: if an act imposes eliminative harming, then it cannot also impose opportunistic harming. The two types of harming cannot be combined in one and the same act. We can call this the *Non-Combinability Claim*. The Non-Combinability Claim entails that, if defensive torture can be established as eliminative, it is not opportunistic.[31]

---

[29] Quinn (1989), p. 344.
[30] Quinn (1989), p. 344.
[31] Kaufman's implicit acceptance of this claim, or something close to it, would explain why he thinks that instrumental violence can, with no additional

The Non-Combinability Claim is questionable. It seems to me that the harming inflicted on individuals in cases of defensive torture admits of a *hybrid* characterization: these victims are harmed in both an eliminative *and* an opportunistic way. To see why, reconsider *Poisoning 4*. In *Poisoning 4*, *Y* is tortured in order to disclose the information that will eliminate the threat he poses. Steinhoff thinks that this feature is enough to qualify the harm as eliminative. I agree. But *Y*'s harming is not *merely* eliminative; it has a more complex character than that. For it cannot be reasonably denied that, in being tortured for information, *Y*'s presence is also used to gain an advantage that, in these circumstances, could not otherwise be gained. *Y* has the antidote, and he is tortured in order to get him to give it to *Z*. That property of the harming inflicted on him suffices to qualify it as opportunistic. (Imagine that *Y* kept slipping out of consciousness, or showed signs that he was about to die. That would frustrate *Z*'s aims. *Z* would therefore take himself to have reason to keep *Y* alive, and keep him awake, in order to continue torturing him in order to get the information to save *X*. How could such treatment fail to be instrumental?) So, in summary, the fact that *Y* poses the initial threat makes *Z*'s harming of him eliminative in character, and the fact that *Z* tortures *Y* for information makes *Z*'s harming of him opportunistic in character. *Y*'s torture is therefore both eliminative and opportunistic.

Because defensive torture is partly opportunistic, it invites a significant comparison with penal torture. Why? Because it would be hard to deny that the explanation of why *Y* can be liable to be harmed in ways which are not simply eliminative is that he has done something he ought not to have done: namely, started an attack on *X*. His hard treatment now is therefore a response to *what he has already done*. Of course, *Y*'s attack on *X* is still *ongoing*, so defensive torture is not a pure case of punishment. But there were two elements of harming contained in acts of defensive torture that we needed to accommodate: eliminative harm and opportunistic harm. The defensive angle is reflected in the presence of eliminative harm. But the presence of opportunistic harm invites us to go beyond defensive considerations, in order to see where the relevant justification might be coming from. It appears to come from a place which looks very like punishment. Penal torture is not discussed in its own right by Steinhoff. This is because, presumably, he does not think it needs to be discussed: penal torture is beyond the pale. If

---

commitments, be inflicted on innocent bystanders as well as those who are liable. Without the Non-Combinability Claim, that inference is more difficult to explain.

defensive torture has properties which make it significantly analogous to penal torture, then moral alarm bells should be ringing.

The second imagined counter-reply is considerably weakened in light of the argument that defensive torture co-instantiates both eliminative and opportunistic harming. Of course, we might continue to prefer torture over death. But that point, *by itself*, establishes very little. I might, in fact, prefer to be sadistically tortured than to be killed, but of course it would not follow that there could be *any* circumstances in which the infliction of sadistic torture on me was justified. My judgment that sadistic torture is a better fate for me than death does not demonstrate, just because there are circumstances in which killing me would be permissible, that there must also be circumstances in which inflicting sadistic torture on me would be permissible. I assume that Steinhoff is not suggesting otherwise. His comparison between torture and death is supposed to be disciplined by the claim that there are no further significant moral *disanalogies* between defensive torture and ordinary defensive violence. That is the work supposedly achieved by the Analogy Claim and Permissibility Claim. But there *is* a disanalogy, as we have seen: defensive torture involves opportunistic harming, not just eliminative harming. Defensive torture is therefore stationed somewhere in the murky territory between ordinary defensive violence and penal torture. That is not where it should be, if considerations about relative harm and comparative permissibility are going to have any chance of demonstrating that defensive torture is morally permissible.[32]

5. Legalizing and Institutionalizing Torture

Many theorists think that, in certain circumstances, torture would be morally permissible. What they resist is the legalization of torture.[33] Steinhoff's position is subtly different. He resists, not the limited legalization of torture, but its institutionalization. This section will investigate whether there is sufficient distance between institutionalization and legalization for Steinhoff's position to be stable. But I want to start with a different puzzle, concerning the distinction between moral permissibility and legal permissibility. Steinhoff is not, in fact, particularly vulnerable to this puzzle, though other writers seem to be.

---

[32] This is just another way of saying that the Content Objection has now been vindicated.

[33] See, for example, McMahan (2008) and Hare (2014).

Return to *Poisoning 4*. Imagine that we are agreed with Steinhoff that it is morally permissible for *Z* to torture *Y*. But now imagine that it is legally impermissible for *Z* to torture *Y*.[34] What should *Z* do? If it really is morally permissible for *Z* to torture *Y*, then it follows that it is morally permissible for *Z* to break the law which forbids him from torturing *Y*. If, by contrast, *Z* is morally required to obey the law which forbids him from torturing *Y*, then it is morally impermissible for *Z* to torture *Y*, and so it was morally impermissible, after all, for *Z* to torture *Y*. Call this the *Moral/Legal Dilemma*.

Now the force of the Moral/Legal Dilemma might be challenged on the following grounds, which I call the *Analogy Defence*:

> *Analogy Defence*
> Surely the Moral/Legal Dilemma establishes much less than it takes itself to have established, since exactly the same dilemma holds for ordinary defensive violence. Take *Poisoning 1*. Clearly, we can establish that it is morally permissible for *X* to kill *Y* to stop *Y* from poisoning *X*. But the law might hold otherwise. (Perhaps legislators think that a law protecting self-defensive action would encourage other types of violence.) If so, then the same Moral/Legal Dilemma will appear: it will be either morally permissible for *X* to kill *Y*, and thus to break the law which makes it impermissible for him to deploy lethal defensive violence, or it will be morally impermissible for *X* to kill *Y*, on the grounds that *X* should not break the law. There is no *special* application of the Moral/Legal Dilemma to *defensive torture in particular*.

The Analogy Defence is unsatisfactory, for the Moral/Legal Dilemma does not hold as forcefully for ordinary defensive violence. Intuitively, a law which prohibited the killings in *Poisoning 1* and *Poisoning 3* would be clearly *morally unjust*.[35] We do not take the same view of defensive torture. One sometimes gets the feeling in this literature that theorists who take defensive torture to be permissible on purely "moral" grounds are actually *relying* upon the law to stop torture from

---

[34] Perhaps the law reflects worries about possible abuses and the incidence of unjustified violence arising from defensive torture.

[35] It is not being denied that the law is entitled to restrict the *forms* of permissible defensive violence: the law can severely restrict gun ownership, for example.

being practised. These theorists would not, I think, take anything like the same deferential attitude to law which restricted ordinary defensive violence.

Now though Steinhoff is not open to this problem, he may still be vulnerable to some other objections. Steinhoff does not provide necessary and sufficient conditions for institutionalization, but what he says about it permits the following generalization: the institutionalization of torture would involve the creation of a bureaucracy specifically concerned with the regulation of torture. The forms of institutionalization might involve the following: the creation of a cadre of specially trained official torturers; the bureaucratic management of torture warrants, where warrants are applied for, scrutinized, and then issued or refused; or the funding of research into torture techniques. People's jobs would come to depend on torture.[36]

Steinhoff does not, in fact, spend too much time on rejecting full-fledged institutionalization. Perhaps he does not have to, for the problems with it are legion: we might expect further entrenchment of it, or torture "creep"; we could expect specific abuses and wrong calls (due to epistemic frailties); and we could, as a result, expect cover-ups and pernicious forms of institutional loyalty and secrecy. We might further expect a worrying coarsening of the general public's moral sensibility; we would rightly worry about the moral calibre of the individuals who applied for jobs in the torture sector; and so on.[37] I agree that all of these problems make any deep institutionalization of torture a very bad idea. The more delicate question to consider is what connections hold between legalization and institutionalization.

Institutionalization clearly comes in degrees. Steinhoff himself distinguishes between *rudimentary* and *full-fledged* institutionalization.[38] The larger the bureaucracy generated by the legalization of torture, the more full-fledged the institutionalization of it is going to be. But even legalization must count as a rudimentary form of institutionalization: those individuals who would be legally empowered to deploy torture would, after all, be institutional actors. The question is whether more minimal forms of institutionalization would be likely to mutate into more severe or far-reaching forms of institutionalization. Steinhoff resists this worry, but on what grounds?

---

[36] Dershowitz (2002) advocates the introduction of such a torture warrant system. His views are criticized by Steinhoff at ET, pp. 61-67.
[37] ET, pp. 61-67. See also Griffin (2010) for a nuanced picture of the various bad effects we might expect from the institutionalization of torture.
[38] ET, p. 67.

One salient consideration for Steinhoff is that *Ticking Bomb* and *Dirty Harry* are very unusual cases. As he memorably puts the point:

> Institutionalizing torture in order to be equipped for such a rare occasion is like stationing a police officer at a lake somewhere in the wilderness because a parachutist might land in it and drown otherwise.[39]

On his view, then, there is simply no need for any further institutional entrenchment beyond mere legalization. The amassed institutional forces, if they existed, would be largely and pointlessly idle. Call this the *Rareness Argument*.

I find the Rareness Argument complacent, for two important reasons. First, and by Steinhoff's own lights, the structure of interrogational torture *just is* defensive. Its defensive, liability-involving structure entails that we do not have to rely on the emergence of cases where only *massive harms* would be averted by the use of torture. There are plenty of individuals in custody, or under surveillance by the security services, who are suspected of involvement in deadly criminal enterprises, in which innocent people are at risk. Should they be tortured for information? Epistemic grounds might be shaky, of course. But that takes us only as far as the second problem.

As we have already noted, mere legalization already constitutes a minimal form of institutionalization. Even if the confident and principled initial expectation was to use torture only very sparingly, and in rare circumstances, the legal permissibility of torture would then give rise to different inflationary worries. First, there would be *Accountability Worries*. Under legalization, there will be a range of cases in which the employment of torture is an *option*. If torture is not used when it might have been used, and things go badly, then these agents will have to produce a justification for why torture was not used. They will be accountable to the following question: why was torture not used when it could have got results? Second, and relatedly, there will be *Bad Call Worries*. In cases where torture is an option, it will sometimes turn out that torture is deployed against those who lacked any relevant information; these cases will reflect the false judgment that torture could be expected to get results. Accountability Worries and Bad Call Worries will arise together, because the bad calls will be explained by a prior anxiety about accountability.

[39] ET, p. 67.

It is not being denied that there is a difference between different parts of the institutionalization slope, or spectrum; there will be a big difference between minimal institutionalization and full-fledged institutionalization. But the danger with being on a slippery slope is that, as soon as you are on it, you are in danger of slipping down it. Accountability Worries and Bad Call Worries, together with cultures of institutional secrecy and loyalty, would then help to impose some downward pressure on the slope, and help to create a metastasizing effect. What would matter, when all was said and done, was that torture was on the legal books, rather than the exact current bureaucratic implications of its being on the books. Torture should not be legitimated in this way.

Another of Steinhoff's reasons for being untroubled by legalization is explored in the following passage:

> Of course, killing and nontorturous injuring in self-defense are not *precisely* the same as torture in self-defense, but no two cases are the same: otherwise they would not be two cases, but one. In order to draw conclusions from one case for the other it suffices that the second case is sufficiently or relevantly similar. And… the differences between self-defensive killing and self-defensive torture are not such that one could not draw conclusions for the case of legalizing self-defensive torture from the fact that the legalization of self-defensive killing does not lead to abuses on such a scale as to make the social costs of allowing self-defensive killing prohibitive. To claim otherwise seems to amount to nothing less than a belief in some kind of magical or addictive properties of torture that other forms of violence, apparently, do not have…[40]

Since this argument is concerned with the further incidence of violence, I will call it the *Further Incidence Argument*.

It is not immediately clear how the Further Incidence Argument is to be interpreted. On one possible interpretation, it states a *moralized* argument, which can be represented as follows:

(a) Legalizing self-defensive violence has not inspired a wave of unjustified violence.

(b) The legalization of self-defensive violence is relevantly morally similar to the legalization of defensive torture.

And so, given (a) and (b), it follows that:

[40] ET, pp. 73-74.

(c) The legalization of defensive torture should not be expected to inspire a wave of unjustified violence.

We can call this variant the *Moralized Further Incidence Argument*. It can be dismissed without much further ado. The challenge to which the Moralized Further Incidence Argument is supposed to be responding, after all, is that the legalization of defensive torture will be indefensible if legalizing it inspires too much unjustified violence. It can be no answer to *that* concern that, since the legalization of defensive torture is relevantly morally similar to the legalization of ordinary defensive violence, the further incidence of unjustified violence is bound to be similar in each case. *If* the respective incidence rates differ, then the analogy will be spoiled. So this particular worry cannot be tackled by refusing to investigate the incidence rates on the grounds that the analogy is already secure. It is incumbent on defenders of the legalization of defensive torture to investigate the respective rates of incidence in order to demonstrate to us that the analogy can be upheld.

Another and much more sensible interpretation of the argument alters the interpretation of (b), so that it is not the moral similarity between ordinary defensive violence and defensive torture which is doing the work, but some other descriptive feature of the analogy:

(a) Legalizing self-defensive violence has not inspired a wave of unjustified violence.
(b*) The legalization of self-defensive violence is relevantly similar to the legalization of defensive torture in respect of its ability to inspire a wave of unjustified violence.

And so, given (a) and (b*), it follows that:

(c) The legalization of defensive torture should not be expected to inspire a wave of unjustified violence.

It is this more plausible variant I shall have in mind when I refer to the Further Incidence Argument.

What is Steinhoff's justification for (b*)? Steinhoff actually has little to say in defence of it, except for his dismissal of the possibility that torture is "addictive".

This is beside the point. The addictiveness of torture is neither here nor there. Whether the respective incidence rates of unjustified violence are different will depend on the motivations and opportunities available to those individuals who are in a position to use violence in these unwarranted ways. The incidence of unjustified violence due to those who are legally permitted to use ordinary defensive violence will depend on the motivations and opportunities open to them when they are in relevantly different circumstances. Similarly, the incidence of unjustified violence due to those who are legally permitted to use defensive torture will depend on the motivations and opportunities open to them when they are in relevantly different circumstances. And there do, potentially, seem to be systematic differences to worry about on this score. Individuals who would be legally empowered to employ defensive torture would *already* be institutional actors; they would be police officers and other security personnel. By contrast, no interesting common institutional identities obtain among individuals—all of us, basically—who are legally empowered to employ ordinary defensive violence. If differences in the incidence of unjustified violence emerge, they will be plausibly explained by these differences between institutionalization and non-institutionalization. Contrary to what the Further Incidence Argument implies, I see no reason to suppose that such differences would prove to be inconsequential.

6. Brutality, Dehumanization, and Legal Archetypes

In conclusion, I want to note a potentially dangerous *psychological* conviction-shifting effect which, if it took root at all, would do so at the legalization stage, and not just the full-fledged institutionalization stage. And I want to connect that line of thought, in at least an embryonic form, to Waldron's argument about legal archetypes.[41] The central danger here is that legalization would lead to the *dehumanization* of prisoners on whom the infliction of defensive torture was an option.

Again, *Poisoning 4* allows us to provide an embryonic sketch of the relevant dangers. Even if *Z*'s attempt to torture the information about the antidote out of *Y* was unsuccessful—imagine that the antidote is not revealed, and *X* dies—it would not follow, on Steinhoff's view, that *Y* had not been liable to be tortured. In other words, we would not make the *success* of *Z*'s attempt a necessary condition of the permissibility of what *Z* did to *Y*. Regardless of whether *Y* provides *Z* with the

[41] See Waldron (2010). Steinhoff discusses Waldron's views at ET, pp. 74-77.

information *Z* seeks, it will have been legitimate for *Z* to torture him. But then *Z*—and indeed the rest of us—may start to think of *Y* as someone who can be permissibly tortured, regardless of whether *Y* actually produces the information that can save *X*. *Y* has done things which are such that *Y* can be tortured with impunity, regardless of what happens next. As a result, *Y*'s captors start to think about *Y in that very way*: they start to regard *Y* as *someone who can be tortured with impunity, regardless of what happens next*. Now of course, according to Steinhoff, it would be a mistake for *Z*, and the rest of us, to think that *Y*'s liability to be tortured could be detached from *Z*'s strenuous commitment to finding the antidote to the poison. But my complaint is not that *Z*'s reasoning would, by Steinhoff's lights, be error-free. The point, rather, is that there is an erroneous but intelligible trajectory of thought which begins with *Y*'s liability to torture in a narrow range of circumstances and ends with *Z*'s comprehensive disregard of *Y*'s moral status. (Anyone who can be tortured with impunity, regardless of what happens next, is morally insignificant.) *Z* may hope for results when he tortures *Y*, but it will not have escaped his notice that *Y* was liable to be tortured even in the absence of those results. *Z* will have to exert some self-discipline to avoid thinking of *Y* as someone who does not count at all. Would we trust *Z* not to succumb to that view of *Y*? Would we trust *ourselves* not to succumb to that view of *Y*?

We can, I think, connect this line of thought to some of Waldron's work in this area. Waldron's central argument against the legal permissibility of torture makes extensive appeal to the idea of a legal *archetype*, which he describes as "a particular provision in a system of norms which has a significance going beyond its immediate normative content, a significance stemming from the fact that it sums up or makes vivid to us the point, purpose, principle, or policy of a whole area of law".[42] As applied to the treatment of prisoners, and the prohibition on torture in particular, the relevant archetype is a *non-brutality* archetype, which can be expressed in the following way:

> Law is not brutal in its operations; law is not savage; law does not rule through abject fear and terror, or by breaking the will of those whom it confronts. If law is forceful or coercive, it gets its way by methods which respect rather than mutilate the dignity and agency of those who are its subjects.[43]

[42] Waldron (2010), p. 228.
[43] Waldron (2010), p. 232. The discussion at pp. 222-260 is generally relevant.

When we interpret anti-torture laws in this way, we can see them as an important barrier against the formation of dehumanizing attitudes towards those in custody.

Steinhoff does not dismiss the idea of archetypes from the outset, but rejects Waldron's reasoning on the grounds of its one-sidedness. Imagine that a jurisdiction that has refused to legalize torture is now confronted with the *Dirty Harry* case. Further suppose, because torture is not used and the child kidnapper is entirely recalcitrant, that the child dies. Steinhoff now presses the charge that the talk of archetypes is unhelpful, because this anti-torture jurisdiction might be just as easily accused of upholding a rival legal archetype: that of *aiding and abetting crime*.[44] So we have two rival archetypes, if we are in the business of ascribing archetypes at all: Waldron's non-brutality archetype, and Steinhoff's aiding-and-abetting archetype. If torture is categorically excluded as an interrogational option, then we will generate material which upholds the aiding-and-abetting archetype, whereas if torture is not categorically excluded, we have material which nourishes the non-brutality archetype. Neither archetype is obviously preferable to the other. The best suggestion, then, Steinhoff's argument implies, is to drop the notion of an archetype, since it is unlikely to achieve much insight into how values must be balanced in such cases.

This blanket dismissal of archetypes strikes me as premature, for upon closer examination the aiding-and-abetting archetype is unconvincing. Imagine a variant of the basic *Dirty Harry* case, which we might call *Extreme Dirty Harry*. To get *T* to reveal the whereabouts of five suffocating children—he would otherwise be entirely resistant to producing such a revelation—the authorities consider torturing *T*'s mother (to whom *T* is close). But the authorities decline to do so, *T* does not volunteer the information, and the children die. It is an awful outcome, but have the authorities aided and abetted *T*'s criminal activity? Why would we choose to represent their actions in this way? They are refusing to torture an innocent person so that children who will otherwise suffocate can be rescued. In general terms, the authorities accept, in advance, certain limitations on what they can do in pursuit of their aim of protecting the innocent. And sometimes, as in *Extreme Dirty Harry*, this sort of scrupulousness will cost innocent lives. Is the authorities' acceptance of such limitations tantamount to *collaboration* with the criminals they are dealing with? Such a view would surely be difficult to impute to them, since so much else of what they do involves the *obstruction* of criminal wrongdoing, rather than the

---

[44] ET, p. 76.

facilitation of it. If the authorities were to be interpreted as aiding and abetting criminals' wrongdoing, there would be the awkward matter of all the arrests, interceptions, and punishment which constituted counter-evidence for the aiding-and-abetting hypothesis. The non-thwarting of criminal wrongdoing in *Extreme Dirty Harry* would have to be somehow balanced with the fact that such wrongdoing was frequently thwarted in other cases.

Clearly this would be an unpromising line of interpretation. The proper archetype to be imputed to the authorities cannot be the aiding-and-abetting archetype across the board. Neither should it be a combination of the obstruction hypothesis (whenever criminal wrongdoing is successfully thwarted) and the aiding-and-abetting hypothesis (whenever criminal wrongdoing is unsuccessfully thwarted). That would also constitute a fantastically uncharitable reading of the basic data.

As I see it, then, Steinhoff's argument does not impugn Waldron's invocation of archetypes. The avoidance of brutality still appears to be a noble and eminently sensible aim for our legislation to enshrine, even if, on occasion, the results go badly for us.

I conclude that Steinhoff's arguments are insufficiently reassuring. Legalization would in all likelihood promote fuller forms of institutionalization, not because torture is addictive, but because legalization would *already be* a form of institutionalization. Moreover, it is both right in itself, and likely to promote better consequences overall, if our legislation aims to uphold the non-brutality archetype. The resulting lesson is perfectly clear: we should not torture prisoners.[45]

References

Dershowitz, A. (2002) *Why Torture Works: Understanding the Threats, Responding to the Challenge*. New Haven, CT: Yale University Press.

Griffin, J. (2010) "What Should We Do about Torture?", in *Ethics and Humanity: Themes from the Philosophy of Jonathan Glover*, ed. N. A. Davis, R. Keshen, and J. McMahan. Oxford: Oxford University Press.

Hare, C. (2014) "Torture–Does Timing Matter?", *Journal of Moral Philosophy* 11: 385-394.

[45] An earlier version of this article was presented at the Applied Ethics seminar in Leeds. I would like to thank various members of that audience for their helpful and thoughtful comments.

Kaufman, W. (2008) “Torture and the ‘Distributive Justice’ Theory of Self-Defense: An Assessment”, *Ethics and International Affairs* 22: 93-115.

Lang, G. (2014) “Why Not Forfeiture?”, in *How We Fight: Ethics in War*, ed. H. Frowe and G. Lang. Oxford: Oxford University Press.

McMahan, J. (2008) “Torture in Principle and in Practice”, *Public Affairs Quarterly* 22: 111-128.

Montague, P. (1989) “The Morality of Self-Defense: A Reply to Wasserman”, *Philosophy & Public Affairs* 18: 81-89.

Montague P. (1981) “Self-Defense and Choosing between Lives”, *Philosophical Studies* 40: 207-219.

Otsuka, M. (1994) “Killing the Innocent in Self-Defense”, *Philosophy & Public Affairs* 23: 74-94.

Quinn, W. (1989) “Actions, Intentions, and Consequences: The Doctrine of Double Effect”, *Philosophy and Public Affairs* 18: 334-351.

Shue, H. (1978) “Torture”, *Philosophy and Public Affairs* 7: 124-143.

Steinhoff, U. (2013) *On the Ethics of Torture*. Albany: SUNY Press.

Thomson, J. J. (1991) “Self-Defense”, *Philosophy & Public Affairs* 20: 283-310.

Waldron, J. (2010) “Torture and Positive Law”, in his *Torture, Terror, and Trade-Offs: Philosophy for the White House*. Oxford: Oxford University Press.