

Unhinged Frames: Assessing Thought Experiments in Normative Political Theory

MATHIAS THALER*

This article develops a framework for assessing thought experiments in normative political theory. It argues that we should distinguish between relevant and irrelevant hypotheticals according to a criterion of modality. Relevant hypotheticals, while far-fetched, construct imaginary cases that are possible for us, here and now. Irrelevant hypotheticals conjure up imaginary cases that are barely conceivable at all. To establish this claim, the article interrogates, via a discussion of Susan Sontag and Judith Butler's accounts of representations of violence, the frames through which hypotheticals construct possible worlds, and concludes that some frames are better than others at sustaining a link with the world as we know it. Frames that disrupt this link can be charged with failing to offer action-guidance.

Keywords: Framing; imagination; judgement; thought experiments; torture.

Readers of contemporary political theory, in particular within Anglo-American academia, will be familiar with thought experiments. They are invoked to bring out 'our intuitions' about ethically controversial issues, such as abortion;¹ they are employed to test assumptions about politically and morally relevant distinctions, such as the one between killing and letting die;² and they are introduced to show the inherent limitations of certain philosophical doctrines, such as hedonism.³ Thought experiments play an especially pronounced role in applied ethics and political theory, where they serve various interrelated functions, such as stipulating counter-examples, creating mechanisms to tease out the implications of certain positions, providing clarification and issuing invitations to re-imagine age-old problems.⁴

Broadly, we may define thought experiments as devices that invite the reader to examine the normative consequences of a hypothetical situation as though 'the particular state of affairs described in the imaginary scenario were actual'.⁵ Thought experiments make present what is

* Politics and International Relations, University of Edinburgh (email: mathias.thaler@ed.ac.uk). Earlier versions of this article were presented at the Political Theory Research Group meeting in Edinburgh (2015), the PSA General Conference in Sheffield (2015) and the APSA General Conference in San Francisco (2015). The author is grateful to the audiences of all these events for their excellent questions. Special thanks are due to Philip Cook, Liz Frazer, Dustin Howes, Kim Hutchings, Moya Lloyd, Mihaela Mihai, Kieran Oberman and Alan Wilson, who have read various/different versions of this article and proposed highly perceptive and helpful feedback. A debt of gratitude is also due to this *Journal's* three referees for suggesting many improvements to the original manuscript. Finally, the author thanks the Editor, Rob Johns, for expertly navigating the article through the review process and for generously offering guidance throughout. The usual disclaimers apply. The research for this article has benefited from a Marie Curie Career Integration Grant (JUDGEPOL).

¹ See most famously Judith Jarvis Thomson's article on abortion, in which she conjures the image of a famous and talented violinist whose life depends on his being plugged into the renal system of another person: Thomson (1971).

² Certainly, the most prominent example is the so-called 'trolley problem': Foot (1967). On the wider impact of the 'trolley problem', which has had an astonishingly steep career, see Edmonds (2013).

³ Robert Nozick's 'experience machine' is a good example for this kind of imaginary case. See: Nozick (1974), pp. 42–5.

⁴ Walsh (2011), pp. 471–2.

⁵ Walsh (2011), p. 469.

absent, and as such they structurally resemble narrative art forms, such as novels or films.⁶ There is nothing new about the introduction of hypotheticals. From Plato's *Republic* to John Rawls's *Theory of Justice*, imaginary cases have preoccupied philosophers a great deal. Whether the proverbial armchair ought to be a privileged spot from which to reason and deliberate about politics has naturally always been contested. But over the past twenty years, a veritable methodological debate around the status of thought experiments has unfolded.⁷

This article intervenes in this general debate, which encompasses various sub-disciplines in philosophy and political science, from metaphysics to International Relations theory, by attempting to parse productive from unproductive hypotheticals.⁸ In focusing on normative political theory, I shall touch on three dimensions of thought experiments, namely 'what they are; how they work; their virtues and vices'.⁹ To tease out some of the account's implications, I will probe a specific case study in more detail – thought experiments involving torture – that epitomizes the conditions under which imaginary cases might either fail or succeed.

My concern with some types of hypothetical builds and expands on a well-established canon of criticism within political theory, which affirms that 'artificial cases make bad ethics'.¹⁰ To this statement, I am adding that *some* artificial cases make *especially* bad ethics *and* politics, by arguing that many hypotheticals violate basic premises that ought to be heeded whenever the imagination is engaged: they radically disrupt a link with reality as we know it, display no concern with human frailty and precariousness and foreclose alternative frames of representation.

While this article's thrust is hence mostly critical, I also gesture towards a more constructive engagement with thought experiments. As a consequence, it treads new territory that followers of Henry Shue's categorical indictment against artificial cases have thus far shunned: in affirming that some hypotheticals can figure prominently in our normative theorizing of complex cases, I reject his wholesale scepticism and outline a novel defence of their usefulness under specific circumstances. The basis of this defence resides in a thus far unexplored examination of the frames that conjurers of hypotheticals require when constructing their thought experiments.

The argument proceeds as follows: I start by outlining the framework within which my account of productive hypotheticals is situated. This framework is based on a democratic view of political theory as a civic activity, which can be contrasted with more scholastic strands that advocate pure speculation as the ultimate goal of normative reasoning. The next section reconstructs the rationale behind the use of hypotheticals in political theory in general and in the recent torture debates more specifically. I shall distinguish between 'realist' and 'clarificationist'

⁶ Due to space constraints, this essay will have little to say about the relationship between thought experiments in the academic literature and narrative art forms. However, I believe that one crucial difference between the two is that novels and films usually tell much richer and more complex stories than the ones constructed in scholarly discourse. On the potential of literature to contribute to, and further advance, moral knowledge, see Carroll (2002).

⁷ The most comprehensive survey of thought experiments that has instigated this methodological debate is Sorensen (1992).

⁸ For a good overview of this debate, including a useful bibliography, see Brown and Fehige (2016). For the use of counterfactuals in International Relations theory, see Lebow (2010); Tetlock and Belkin (1996). For two defences of thought experiments in the realm of applied ethics and political theory, see Mišćević (2013), Walsh (2011).

⁹ See Sorensen (1992), p. 3. Several authors have tried to explore where and how thought experiments go wrong. This article seeks to build on their findings and apply them to the discussion around torture. See Brendel (2004), Walsh (2011).

¹⁰ Shue (1978), p. 141. Aside from Shue's seminal paper from the 1970s, see in particular: Bufacchi and Arrigo (2006); Luban (2009); Shue (2006).

defenders of thought experiments and sketch their rationales for justifying imaginary cases. I then scrutinize Susan Sontag's and Judith Butler's exploration of the linkages between framing, the intelligibility of human suffering and ethical responsiveness. The main lesson to draw from their writings is that frames, far from being innocuous, are always the contingent products of specific decisions about the field of representability.¹¹ The penultimate section utilizes this basic, yet consequential insight to mark the distinction between 'this-worldly' and 'other-worldly' hypotheticals. I will examine two thought experiments about torture (Innocent Jenny as well as the ticking-bomb scenario) and weigh their virtues and vices. Finally, the last section elaborates on three objections and offers a conclusion.

POLITICAL THEORY AND ITS AUDIENCE: A FRAMEWORK FOR THINKING THROUGH HYPOTHETICALS

The goal of this section is to propose a primer for thinking through imaginary cases in political theory. The debate around the proper use of imaginary cases is implicated in discussions around 'how to do' political theory.¹² The reason for this is the following: political theory can, with some justification, be exercised without much direct concern for the real world, from the armchair, as it were. But, to stick with the metaphor, the ultimate purpose of sedentary political theory, its *telos*, surely is to get the audience moving, to elicit certain (positive or negative) reactions from them, to make them perform actions they would otherwise not do.¹³ This is true even if the political theorists themselves, wishing the audience to move in a particular direction, stand (or sit) still. In other words, although it would be mistaken to insist that political theorists must always be directly concerned with action-guidance, we still ought to reject the notion that pure speculation is an option for political theory.

When suggesting that political theory ought to speak to an audience outside academe and develop arguments that address wider issues in the public debate, I do not mean this to be translated crudely into pontificating to those who have not yet been exposed to normative analysis. Such a perspective would have the unfortunate and absurd consequence that political theorists should be held accountable in situations where the public completely ignores them, which is, arguably, what happens most of the time. It would furthermore have the upshot of symbolically elevating political theorists above their audiences, an unattractively elitist image that needs dismantling.

What, then, does it mean for political theory to speak to an audience in the sense described above? The idea is that, since political theorists are concerned with practical as opposed to

¹¹ To avoid a misunderstanding from the start: I refer here to 'framing' in a sense that is distinct from how the term is employed in behavioural economics or cognitive psychology. 'Framing', throughout this article, means the ways in which certain salient facts are represented in a thought experiment.

¹² I am referring here to the debate around the relationship political theory ought to entertain with the real world. The debate is usually organized around the correct interpretation of Rawls's 'ideal/non-ideal' distinction, or around the nascent features of realist political theory. Since I fear that grappling with this literature in more detail would distract me from the overall ambition of this article, I shall be obliged to only hint at it *en passant*. For some notable publications, see: Farrelly (2007); Geuss (2008); Gilabert (2011); Mills (2005); Valentini (2012).

¹³ My understanding of political theory as speaking to an audience has been inspired by Bernard Williams's reflections on the 'listeners' to whom political theorizing is addressed. With Williams, I believe that political theory can only fulfil its vocation of being heard if its diagnoses and prescriptions somehow link up with the self-interpretations of those who have been implicitly or explicitly spoken to. To complicate matters, however, Williams uses the term 'audience' in a more technical sense for those who will be likely *actually* to read the text. While this distinction between 'listeners' and 'audience' might perhaps make sense in some regards, it is not substantial to my argument. See Williams (2005), pp. 56–60.

theoretical reasoning, their work is best understood as contributing, in various complex and often indirect ways,¹⁴ to discussions ‘with their fellow citizens as equals’.¹⁵ Yet another way of expressing this thought is to describe political theory as ‘principled social criticism’.¹⁶ Evidently, this does not entail that such a democratic orientation towards the public debate would be the only criterion by which we may assess political theory’s propositions; but it provides us with a standpoint from which we are able to perceive more clearly when and how political theorists fail in taking practical reasoning seriously.

This is not an uncontroversial view. Amongst others, G. A. Cohen famously defended the speculative nature of political theory in asserting that the question ‘is not what we should do but what we should think, even when what we should think makes no practical difference’.¹⁷ This article does not directly contribute to the intricate debate around political theory’s standing *vis-à-vis* the real world (often couched in terms of the ideal/non-ideal distinction or in terms of calls for more ‘realism’ in political theory). Suffice it here to state that, within the framework informing this article, a strict separation between thinking and acting, as suggested by Cohen, is not a desirable aspiration for political theory.

To apply this point to our topic: once the driving force behind political theory becomes ‘what we should think’, rather than ‘what we should do’, the very purpose of thought experiments shifts from action-guidance to abstract ruminations. Consequently, the debate in political theory would run the risk of becoming overly self-referential and scholastic, detached from the wider issues that animate the public debate. Promoting pure speculation is hence problematic and should be resisted, especially in cases where the use of violence is discussed.¹⁸

The article’s concrete proposal entails, in a nutshell, that conjurers of hypotheticals should primarily strive to enhance the audience’s capacity to judge difficult situations in the real world. Productive thought experiments engage the imagination in such a way as to enable these judgements. In that sense, they can be deemed relevant for our current predicament. When hypotheticals fail, they fail for different reasons. I shall distinguish between two archetypical instances: (1) failure to develop ‘imaginary grip’, i.e. the incapacity of a given hypothetical to engage the imagination of the reader or viewer; (2) failure to track certain key features of the real world, which must not be forfeited in imaginary cases, i.e. the incapacity of a given hypothetical to account for what is possible for us, here and now. Further to explain the heuristic terminology used in this article, I shall call productive thought experiments ‘this-worldly’, while reserving the term ‘other-worldly’ for hypotheticals that do not manage to improve the audience’s capacity to judge real-world dilemmas.¹⁹ The vices and virtues of concrete thought experiments crucially depend, as I shall demonstrate in the following, on the frames that political theorists draw on when devising imaginary cases.

¹⁴ The compound phrase ‘complex and often indirect’ is crucial for my interpretation of how political theory ought to speak to its audience. Since I do not advocate pontificating to the public as a desirable outlook for political theorists, I merely seek to reject a vision of political theory that is radically opposed to action-guidance altogether. In the next paragraph, I call this the ‘pure speculation’ strand in normative political theory.

¹⁵ Tully (2008), p. 8. On this point, see especially Laden (2011).

¹⁶ Shapiro (2005), p. 173.

¹⁷ Cohen (2008), p. 268.

¹⁸ For a biting critique of scholastic reason more widely, see Bourdieu (2000).

¹⁹ The terminology mirrors, in some respects, Jakob Elster’s distinction between realistic and outlandish cases in ethics. While I find his argument largely persuasive, I reject his conclusion that the criterion of action-guidance cannot be enlisted for assessing thought experiments; see Elster (2011).

OBSERVING TORTURE FROM THE ARMCHAIR: 'IMAGINE A SITUATION IN WHICH ...'

The previous section established the conceptual platform from which the interpretation of hypotheticals will be launched. This section explains how thought experiments are commonly defended in contemporary political theory. As any reader of this literature knows, thought experiments are frequently used to fulfil a number of different functions, from illustrating abstract arguments to formulating counter-examples. The great variety of imaginary cases within the contemporary debate testifies to their vital importance for a discourse that stands at some critical distance to the real world. The use of hypotheticals plays an important part in this process for it allows the political theorist to perceive and present things from a new, illuminating angle.

Beginning a conversation with a phrase such as 'imagine a situation in which ...' opens up a discursive space liberated from the conventions on which our comprehension of the real world rests; and yet, whenever hypotheticals are introduced in normative political theory, the question naturally arises how they relate to the real world – what their precise purpose is supposed to be. The reason for this worry is simple. The more we distance ourselves from the world as we know it, the more pressing becomes the imperative to clarify the imaginary case's connection with reality. In other words, if one of the main goals of hypotheticals is to enable deliberations about complex cases, then we need to ensure that the thought experiment is sufficiently in tune with the reality to which it speaks, lest it becomes irrelevant for the specific situation under scrutiny.

Imaginary cases involving torture are exceptionally well suited to disclose this tension inherent in political theory's engagement with hypotheticals. The most commonly invoked hypothetical in discussions around torture has been the so-called 'ticking-bomb scenario'.²⁰ There are several descriptions of this hypothetical, but they all share a roughly similar design: '(1) the lives of a large number of innocent civilians are in danger; (2) the catastrophe is imminent, therefore time is of the essence; (3) a terrorist has been captured who holds information that could prevent the catastrophe from occurring'.²¹ The ticking-bomb scenario refers, in short, to a situation in which the absolute prohibition of torture ought to be suspended to save innocent lives.

While the scholarly literature has dealt extensively with the thought experiment, we still lack a clear sense of why political theorists draw on the ticking-bomb scenario, that is, what kinds of reason they invoke to justify their use of it and similar hypotheticals about torture. I shall suggest that defenders of this imaginary case typically come in two varieties. On the one hand, there are 'realists'²² whose appeal to the thought experiment is intended as an attack on the absolute ban on torture in law and public policy. Alan Dershowitz, for example, enthusiastically advocates the issuing of 'torture warrants' by independent judges, because he thinks that in exceptional situations interrogational torture would be practised anyhow. On this account, the ticking-bomb scenario is not all that far-fetched – it merely dramatizes certain key features of reality to bring out presumably widely held ideas about the problem of dirty hands. For realists, the thought experiment presents us with an 'extraordinarily rare', but 'real and recurring'²³

²⁰ For an excellent discussion of the emergence of the ticking-bomb scenario as the focal point of discussions around torture, see Farrell (2013).

²¹ Bufacchi and Arrigo (2006), p. 358.

²² In using the term 'realist' to characterize Dershowitz's position, I do not mean to suggest that he indeed accurately captures the reality of torture. On the contrary, I agree with critics such as Bob Brecher, who have conclusively shown that Dershowitz pays astonishing little attention to how torture works. The moniker 'realist' rather signifies that, on this reading, the ticking bomb scenario is intended to depict a stylized version of reality. See Brecher (2007), pp. 14–39.

²³ Dershowitz (2002), p. 140.

situation; far from being a figment of our imagination, it tells us something crucial about the way the world actually looks, albeit in especially stark colours. Oren Gross, who otherwise objects to Dershowitz's endorsement of torture warrants, concurs with the proposition that hypotheticals like the ticking-bomb scenario are 'real, albeit rare. Ignoring them completely, by rhetorically relegating them to the level of "artificial", is utopian or naïve, at best'.²⁴

Contrast this with what one may call the 'clarificationist' camp. Clarificationists emphasize that the ticking-bomb scenario is a powerful device because it functions as a counterfactual screen against which the moral permissibility of torture can be explored. On the clarificationist view, hypotheticals involving torture must not be mistaken, under any circumstances, for statements about the real world. Fritz Allhoff makes this point in his defence of torture when he stresses that '[m]oral theory is logically distinct from moral practice'.²⁵ The underlying image of theorizing is one of complete disconnection from reality. In sharp contrast with authors like Dershowitz or Gross, Allhoff and others are thus agnostic as to whether there are actual situations that might somehow resemble the set-up of the ticking-bomb scenario. Allhoff emphasizes that normative theorizing about torture should not, and indeed must not, be conflated with empirical observations about how torture works or when it should actually be allowed and implemented.²⁶

Comparing the realist and the clarificationist interpretations, it becomes clear that those who believe the ticking-bomb scenario tells us something about the real world employ the hypothetical for radically different reasons than those who wish to investigate the moral permissibility of torture via the introduction of a thought experiment. Realists refer to thought experiments so as to throw a real and recurring, albeit rare, situation into sharper relief; clarificationists reject the notion that hypotheticals, such as the ticking-bomb scenario, reflect elements of the world as we know it, and insist on the counterfactual nature of their inquiry. In each of these camps, the imagination is engaged for different reasons: the former use the thought experiment to sharpen a specific representation of reality, the latter appeal to our sense of the possible to shine an unconventional light on a complex case.

If one were to ponder a critical response to each of these defences of imaginary cases, different routes would have to be chosen. A potentially effective rejoinder to the realist position, which will be further explored in this article's penultimate section, would be to propose that the hypothetical fundamentally misconstrues the reality upon which it pretends to rely. Instead of approximating a 'real and recurring' situation, the critic could point out that the ticking-bomb scenario in fact distorts crucial features of the world as we know it. However, this response will obviously not work when dealing with clarificationists – they willingly acknowledge that reality is much messier than their imaginary cases, which is precisely why it is much harder to formulate a cogent answer to the clarificationist camp.

Clarificationist Strategies in Defending Hypotheticals: Three Paradigmatic Positions

Amongst the most outspoken clarificationist defenders of thought experiments about torture, are Frances Kamm,²⁷ Uwe Steinhoff²⁸ and Jeff McMahan.²⁹ In singling out these three authors,

²⁴ Gross (2004), p. 234.

²⁵ Allhoff (2005), p. 247.

²⁶ For those inclined to be swayed by the standard terminology in normative political theory, in Rawlsian terms, Allhoff's move could be associated with asserting the priority of ideal over non-ideal theory. See Rawls (1999), pp. 215–18.

²⁷ For our purpose, see especially Kamm (2008).

²⁸ Steinhoff 2013.

²⁹ McMahan, 2008.

I intend to reconstruct the conceptual background within which thought experiments typically occupy a central place. The point of this reconstruction is not to delve into the minutiae of a few imaginary cases, but rather to delineate the rationale behind their invocation.

Kamm's method aims at the production of hypotheticals whose primary purpose is to test moral intuitions. Her main supposition is that these intuitions remain stable even when imaginary cases become extremely detached from reality. This 'technique of *equalizing cases*'³⁰ involves modifying imagined scenarios such that morally salient factors can be treated discretely and in isolation from each other. In conjuring up thought experiments, Kamm thus grapples with intuitions as if they were 'fixed data points in moral reflection'.³¹ The upshot of this method is the design of hypotheticals that certainly stretch the limits of what some would consider imaginable.³² Kamm introduces a number of thought experiments to unpack the permissibility of torture. Whereas she engages with existing definitions of torture, her main interest is to use moral intuitions to explore what torture is, philosophically speaking.³³

Uwe Steinhoff's use of imaginary cases resembles Kamm's. He, too, employs hypotheticals to question the absolute ban on torture. Steinhoff suggests that torture is justifiable (and not only excusable) on the grounds of self-defence. Since people have a right to defend themselves against aggression, this right covers both lethal force and torture. Steinhoff's is a rights-based defence of torture. Torturing somebody is not only permissible in 'extreme emergencies', when deontological constraints on what we ought to do are overridden by utilitarian concerns for saving the innocent in large numbers; it is permissible whenever innocent lives are endangered, independent of their numbers. His conclusion is neatly summarized in the following paragraph:

Since people have a right even to *kill* a culpable aggressor if, in the circumstances, this is a proportionate and necessary means of self-defence against an imminent threat, and since most forms of torture are not as bad as killing, then people must also have a right to torture a culpable aggressor if this, too, in the circumstances, is a proportionate and necessary means of self-defence against an imminent threat.³⁴

In order to flesh out this point, Steinhoff draws on a number of thought experiments that aim to disprove the position of absolutists. This strategy involves the construction of hypotheticals such as Innocent Jenny, who engages in what Steinhoff calls 'justified self-defensive rape'³⁵ and about whom we will hear more at a later stage. The purpose of these examples is unmistakable: since absolutists lack the imagination to consider cases that challenge their

³⁰ Kamm (2008), p. 427.

³¹ Brand-Ballard 2007.

³² Consider the following hypotheticals (taken from Kamm (2006), pp. 42, 52, 55), which are all supposed to unearth distinctions relevant to terrorism: Suppose that a pilot mistakenly believes that the people on the ground are cows. He has no intention of harming and terrorizing people but only of bombing a building to harm some cows in order to terrorize other cows into trampling a munitions site. ... We bomb some trees because we know that people will think we are trying to kill them and they then become terrorized and pressure for a change of policy. However, we do not and would not harm anyone. ... [C]onsider releasing a flock of butterflies over an opponent's population when we know they are very (irrationally) frightened of butterflies. This involves taking advantage of their irrationality and so is demeaning in a way that reasonable and truth-produced terror is not. Trying to get people to behave in a cowardly way is also preying on their imperfections.

³³ One of the main tools she employs is the trolley case, which plays a crucial role in the investigations of all clarificationists. See Kamm (2008), p. 89.

³⁴ Steinhoff (2010). Henry Shue influentially criticized the idea that killing might be worse than torturing in his seminal paper on torture. See Shue (1978), pp. 129–30.

³⁵ Steinhoff (2013), p. 149.

seemingly steadfast commitment to the torture prohibition, we need extreme hypotheticals to awaken them from their dogmatic slumber.

Finally, Jeff McMahan has not only contributed to the reinvigoration of analytical Just War theory,³⁶ he has recently also reflected on methodological issues arising from the torture debates. McMahan concurs with Steinhoff that torture cannot be absolutely opposed. Since there is a moral equivalence between killing and torturing (in self-defence), it should be acknowledged that torturing the culpable to defend the innocent is permissible.³⁷ Nevertheless, McMahan does not endorse the legalization of torture. *Pace* Dershowitz, he is adamant about the requirement to separate moral permissibility from legal authorization. McMahan hence agrees with those who reject attempts to institutionalize torture in the real world. Yet, the reason why we should oppose torture warrants is not because torturing as such is morally wrong; it is rather that there is a high risk innocent people would be wrongly tortured if warrants were in fact issued.³⁸

What are hypotheticals for, then? Their purpose, according to McMahan, is to ‘filter out irrelevant details that can distract or confuse our intuitions, thereby allowing us to focus on precisely those considerations that we wish to test for moral significance’.³⁹ Thought experiments clean up the messiness of the real world in order to lay bare what the right course of action would ideally be. What is more, although he does not proffer any empirical evidence for this belief, McMahan maintains that imaginary cases have the potential advantage of generating consensus – perhaps even intercultural agreement – as regards intuitions about morality. References to historical cases, which have been notably present in much of the recent literature on Just War theory, will inevitably be tainted by superfluous facts that disorient the moral compass.⁴⁰

While there are, of course, considerable differences between Kamm’s, Steinhoff’s and McMahan’s positions, they subscribe to a common rationale for the use of thought experiments and share family resemblances not to be found in the ‘realist’ camp. To characterize these family resemblances, we may observe that the following argumentative steps are taken typically to defend hypotheticals about torture.⁴¹

1. Since it is an open question whether torture should be absolutely prohibited or not, one must interrogate the ban with as much care as possible. This requires a ‘bracketing’ of the taboo-like prohibition, so as to inquire into the moral permissibility of torture.
2. One way of establishing the moral permissibility of torture is to refer to intuitions about other areas of human interaction in which the use of force might reasonably be justified, such as self-defence or the defence of innocent others. It is assumed that various acts of violence (killing, raping, torturing) can be subjected to the same kind of normative reasoning.
3. Thought experiments can then be designed to test those intuitions by eliminating distorting factors in reality. In so doing, they equalize otherwise morally over-determined cases and thereby make them comparable.

³⁶ McMahan 2002; McMahan 2009.

³⁷ McMahan (forthcoming), pp. 6–8.

³⁸ McMahan (forthcoming), p. 11.

³⁹ McMahan (forthcoming), p. 3.

⁴⁰ Most famously, Michael Walzer can be credited for creatively combining the interpretation of historical cases with normative reasoning around warfare. Revisionist Just War theorists, like McMahan, position their whole project in opposition to Walzer’s; see Walzer (2006).

⁴¹ To be sure, I am not claiming here that all clarificationists explicitly subscribe to the same positions. Rather, my goal is to distil, from a variety of authors, the ideal-typical rationale behind a specific defence of thought experiments. This heuristic exercise seeks to condense the positions of clarificationists into a coherent view.

4. It is normatively irrelevant how far-fetched and outlandish thought experiments become – intuitions will remain stable across the spectrum between standard and extreme hypotheticals.

Perhaps the most obvious respect in which this chain of arguments could be broken is by questioning the quasi-naïve appeal to intuitions. Kamm in particular has been challenged for her heavy reliance on intuitions to determine the meaning of torture.⁴² Yet, intuitions, the objection goes, are not simply data points comparable to scientifically observable facts – they express particular views that are path-dependent on cultural backgrounds and social trajectories. In this respect, it is rather telling that no clarificationist defender of thought experiments about torture has ever seriously tried to find out whether a ‘large majority of people from a variety of cultures’⁴³ would actually subscribe to the same intuitions regarding violence as an Oxford don or Harvard professor.

This lacuna has been recognized as a major shortcoming in contemporary analytical philosophy. Polemically, thought experiments have been labelled ‘intuition pumps’: ‘fiendishly clever devices’ that ‘cajole you into declaring your gut intuition without giving you a good reason for it’.⁴⁴ The burgeoning field of experimental philosophy has emerged specifically to compensate for the inadequacies of such an unexamined appeal to intuitions, by incorporating findings in psychology and by taking ordinary people’s views seriously.⁴⁵ According to experimental philosophers, it might perhaps be possible to establish a commonly shared stance on torture, among ordinary people with diverse cultural backgrounds and social trajectories, but such a claim would have to be empirically substantiated, rather than simply postulated and ‘proven’ through hypotheticals.⁴⁶ I shall not continue this line of criticism. Rather, my subsequent inquiry hones in on modality as the main criterion for assessing thought experiments. Before that, however, the next section casts a look at the often-ignored effects of framing.

THE EFFECT OF FRAMES ON THE INTELLIGIBILITY OF SUFFERING

In this section, I steer our attention to two authors who have grappled in depth with representations of violence: Susan Sontag and Judith Butler. My argument here is that Sontag’s and Butler’s observations on the linkages between framing, the intelligibility of suffering and ethical responsiveness can be fruitfully utilized to elucidate imaginary cases, too. This applies not only to the debates around torture, but also to thought experiments within normative political theory more generally.

Sontag develops her account in a long essay on war photography and later in a newspaper article on the human rights violations in the Iraqi Abu Ghraib prison.⁴⁷ In both texts, she is intrigued by the capacity of photography and representations of violence more generally to

⁴² Wisniewski (2014), pp. 658–9.

⁴³ McMahan (forthcoming), p. 3.

⁴⁴ Dennett (1991), pp. 282, 397.

⁴⁵ For a good overview of the field in general, see Knobe and Nichols (2008). On experimental philosophy and moral reflection, see Alfano and Loeb (2014).

⁴⁶ To my knowledge, there is only one paper in experimental philosophy that tangentially touches on the torture prohibition, in the context of an analysis of ordinary people’s views on intentionality, i.e. Cova and Naar (2011). There is also a deeper worry about the politics of appealing to intuitions: that they simply cover up traces of power and privilege that need to be scrutinized critically. For an insightful discussion of the problematic use of intuitions about justice, see Geuss (2008), pp. 101–3.

⁴⁷ Sontag 2003; Sontag 2004. These texts are based on Sontag’s earlier work on photography: Sontag (1977).

either prompt ethical responses or foreclose them.⁴⁸ Sontag begins by suggesting that a ‘photographic image ... cannot be simply a transparency of something that happened. It is always the image that someone chose; to photograph is to frame, and to frame is to exclude’.⁴⁹

The exclusionary act of framing is important for how we perceive wartime photography. While it might have been the case in earlier days that wartime photography had a cautionary and shocking effect on the audience – during the First World War, the German photographer Karl Friedrich was stopped by state censors from publishing his antimilitarist book *Krieg dem Kriege!*, with its gruesome images of mutilated bodies and wounds to the face, for fear of debilitating the national war effort – today, the situation looks patently different. Since conflict reporting turned over the past decades into a global business for news outlets, ‘ideologues of photography have become increasingly concerned with the issues of exploitation of sentiment (pity, compassion, indignation) in war photography and of rote ways of provoking feeling’.⁵⁰ The fact that various emotions can be triggered through images, that the consumers of reportages are liable to be manipulated by embedded reporters, implies that war photography is radically open to abuse. When watching pictures of atrocity, different sensitivities can be cultivated: ‘To steel oneself against weakness. To make oneself more numb. To acknowledge the existence of the incorrigible’.⁵¹

The chief lesson to take from Sontag is that the frames of violence are inherently unstable and open to contestation: under favourable circumstances, they can render distant suffering intelligible; under unfavourable conditions, visual representations of violence simply satisfy, and even exacerbate, voyeurism. War photography, or indeed any representation of violence, is never only about documenting historical events. It matters who takes the pictures, who makes them available, where they are published, and how we look at them.⁵²

Mitigating the Exclusionary Effects of Frames

In *Precarious Lives* and *Frames of War*, Judith Butler elaborates on some of themes studied by Sontag. Her interest lies with the effects of framing on the intelligibility of suffering:

My point ... is to suggest that, whether and how we respond to the suffering of others, how we formulate moral criticisms, how we articulate political analyses, depend upon a certain field of perceptible reality already being established. This field of perceptible reality is one in which the notion of the recognizable human is formed and maintained over and against what cannot be named or regarded as the human, a figure of the nonhuman that holds the place of the human in its unrecognizability.⁵³

This discussion of framing partakes in a larger project of interrogating various implications of the War on Terror. Butler scrutinizes the ways in which human life can become grievable: when does a human being count as worthy of sorrow and empathy? During the American wars

⁴⁸ Although Sontag does not reference them, social scientists have been examining for some time the relationship between distance and suffering critically against the backdrop of a transformed media landscape. See, for example, Boltanski (1999); Cohen (2001).

⁴⁹ Sontag (2003), p. 46.

⁵⁰ Sontag (2003), p. 80.

⁵¹ Sontag (2003), p. 98.

⁵² This uncontroversial observation is essential because Sontag maintains that photography is not the right medium for enhancing our understanding of violence at all. But images are exceptionally useful for remembering trauma: ‘Harrowing photographs do not inevitably lose their power to shock. But they are not much help if the task is to understand. Narratives can make us understand. Photographs do something else: they haunt us’ (Sontag (2003), p. 89).

⁵³ Butler (2009), p. 64.

in Iraq and Afghanistan, this question has evidently been pressing. Butler surmises that some lives, those of the vanquished and oppressed, simply do not appear grievable – at least to the Western mind. There is a ‘hierarchy of grief’⁵⁴ that makes it impossible to recognize those lives as worthy of sorrow and empathy. While Western lives, and in particular those of the victims of the 9/11 attacks as well as the soldiers fallen in Iraq and Afghanistan, have been mourned on the front pages of widely read newspapers in the United States, Butler observes that ‘there are no obituaries for the war casualties that the United States inflicts, and there cannot be. If there were to be an obituary, there would have had to have been a life, a life worth noting, a life worth valuing and preserving, a life that qualifies for recognition’.⁵⁵

As one commentator has shown, this statement is exaggerated.⁵⁶ Ever since the wars have started, people in the affected countries have of course mourned the deaths of their loved ones and compatriots. Butler wants to make a point about the field of representability in which violence is enacted, perceived and memorialized. Without symbolically disregarding the lives of the vast majority of victims in Iraq and Afghanistan, it would have been impossible to embark on the mission. In her analysis of torture and photography, concentrating in particular on the photos taken in the Abu Ghraib prison, Butler explores this thought further:

[We] cannot understand this field of representability simply by examining its explicit contents, since it is constituted fundamentally by what is cast out and maintained outside the frame within which representations appear. We can think of the frame, then, as active, as jettisoning and presenting, and as doing both at once, in silence, without a visible sign of its operation and yet effectively. ... Prior to the events and actions that are represented within the frame, there is an active, if unmarked, delimitation of the field itself, and so a set of contents and perspectives that are not shown, never shown, impermissible to show.⁵⁷

If such delimitation predates the representation of violence, we may ask whether there are standards against which the legitimacy of delimiting the representable might be measured and compared. Admittedly, Butler herself is not interested in this kind of normative probing;⁵⁸ yet, she still seems to be committed to the notion that not all forms of framing are equally innocuous.⁵⁹ If the existence of a hierarchy of grief cannot be denied, we should perhaps attempt to examine how steep the hierarchy is and how we may eventually flatten it. This implies asking whether all non-Western life is barred from the realm of compassionate engagement, or whether in exceptional circumstances ungrivable life can break through the frames initially set. Indeed, Butler highlights the malleability and contestability of what can be represented:

There are ways of framing that will bring the human in its frailty and precariousness into view, allow us to stand for the value and dignity of human life, to react with outrage when lives are degraded or eviscerated without regard for their value as lives. And then there are frames that

⁵⁴ Butler (2004), p. 32.

⁵⁵ Butler (2004), p. 34.

⁵⁶ Zehfuss 2009.

⁵⁷ Butler (2007), p. 953.

⁵⁸ Butler’s methods of choice are deconstruction and genealogy. On her specific interpretation of genealogy, see Stone (2005).

⁵⁹ My reading of Butler as supplying the basis for a normative analysis of framing might be criticized for misrepresenting her understanding of norms. It has indeed been demonstrated that Butler is most interested in exploring the nexus between norms and violence, or, in what Terrell Carver and Samuel Chambers call ‘normative violence’ (2008, pp. 75–91). Butler’s interpretation of torture photography traces the link between norms and violence as well.

foreclose responsiveness, to be understood as the negative action of existing frames, so that no alternative frames can exist; for them to exist and to permit another kind of content would perhaps communicate a suffering that might lead to an alteration of our political assessment of the current war.⁶⁰

The lesson of Butler's approach in general is that novel frames are needed to oppose those that make responding to distant suffering impossible. Naturally, neither Sontag nor Butler suggests that frames, which are sensitive to the suffering of victims of war and torture, on their own would be adequate to move spectators to act; much more has to change for violence to subside. Still, the idea of bringing the 'human in its frailty and precariousness into view' provides a good starting point for rethinking the ways in which representations might be transformed so as to mitigate the inevitably exclusionary effects of framing.

On the reading of Butler promoted here, the point about foregrounding the precariousness and frailty of human life is not to categorically deny that some lives are strong and secure, even in extremely adverse conditions. After all, we possess ample evidence for heroic resistance to suffering. Rather, Butler reminds us that such strong and secure lives are altogether exceptionally rare, and that representations of violence that prioritize them should be looked at with much scepticism. Ordinary lives, the argument goes, are characterized by the distinctive features of precariousness and frailty.

Butler's position *vis-à-vis* the field of representability possesses vital implications for our topic. Depictions of torture that steer the gaze away from the frail and precarious existence of those on the rack not only symbolically harm the tortured, they also perniciously manipulate the viewers in so far as they project an incomplete and one-sided picture of reality in which the value of some lives is discounted against the value of others. By suppressing the humanity of torture victims, such representations, imagined or not, perpetuate and entrench political agendas through which outrage at violence can be kept at a bare minimum. What is more, some depictions of torture quite literally crowd out alternative frames – the ticking-bomb scenario might be a good example of this phenomenon. A multiplicity of frames would allow the interrogation of hegemonic representations of violence. That is why Butler is adamant about the dire need for many frames.

In sum, responsiveness to the plight of others, no matter who they are or what they have done, is one type of taking action in the face of violence. In this sense, responding in ethically appropriate ways to representations of violence means literally standing up for humanity and resisting dehumanizing policies.

MODALITY MATTERS: AN ICONOGRAPHY OF THIS-WORLDLY AND OTHER-WORLDLY HYPOTHETICALS

This section asks how Sontag and Butler may help us reflect critically on imaginary cases and identify a yardstick to hold apart this-worldly from other-worldly hypotheticals. Recall how those whom I have dubbed 'clarificationist' defenders of thought experiments maintain that contemporary or historical examples taken from the real world are inimical to fine-grained moral appraisal because they are suffused with scattered details, some of which deserve attention whilst others do not. Real-world cases mix up the normatively salient with the normatively extraneous. Imaginary cases, however, have the benefit of stripping reality of all superfluous information so that intuitions can be tested in an environment free of interference: less noise, more sound.

⁶⁰ Butler (2009), p. 77.

I have already gestured towards one respect in which this idea can be dismantled – by shattering the myth of intuitions as simply given, or as stable even at the fringes of the imaginable, or as widely shared amongst people from different cultural traditions. Apart from this route, is there another way in which thought experiments could be critically appraised? Sontag and Butler open up an alternative path of critique, which has so far remained unexplored in the literature. According to their criticism, frames are not neutral devices that simply serve the purpose of rendering reality representable; they are the products of decisions that can be subjected to critical scrutiny.⁶¹

To be sure, thought experiments do not pretend to represent reality; they summon us to imagine a moral universe that is different from the one we currently inhabit. But in order to do so, they still need to latch onto something in the mental space of the readers so as to activate their intuitions, feelings and reasons. A thought experiment that does not occupy this mental space will lack ‘imaginative grip’.⁶²

One way in which hypotheticals might fail would hence be if they invited the audience to envision a situation that exceeds their imaginative powers. In one sense, such a case would barely qualify as a proper hypothetical: its incapacity to move the reader would prevent it from taking off at all. Naturally, it remains empirically open to debate whether some of the scenarios conjured up by Kamm, Steinhoff and McMahan manage to develop imaginative grip or not. It simply depends on the readers’ habits and sensibilities whether they will be able to accommodate certain extreme hypotheticals in the repertoire of their mental space. However, I will assume here that we can accept the invitation to think about torture the way Kamm, Steinhoff and McMahan want us to.

Is there another way in which we could say that some imaginary cases fail to achieve what they have set out to do? Perhaps they move the audience, but in the wrong direction.⁶³ In the introduction, I stipulated we draw a heuristic distinction between ‘this-worldly’ and ‘other-worldly’ thought experiments. Here is how this distinction could be articulated. Assessing thought experiments requires us to look into what kinds of possible world they construct. In unpacking this issue, we may distinguish two types of modality underpinning thought experiments: ‘possible *simpliciter*’ and ‘possible for us, here and now’.⁶⁴ In the following, I explicate in more detail how to separate these two kinds of modality, but I shall start here by underlining the *relevance* of imaginary cases. Thought experiments that are possible for us, here and now, are relevant in the specific sense that they speak to the world we currently inhabit – they address an audience that goes beyond the boundaries of academe and endeavour to facilitate the judgement of complex cases. In order to remain relevant, these hypotheticals need

⁶¹ For example, although she remained outside the frame, the woman who took the photos in the Abu Ghraib prison, Specialist Sabrina Harman, clearly played a role in the depicted reality – a fact for which she was later sentenced to a prison sentence and discharged from the United States Army for bad conduct. See Gourevitch 2008; Morris 2008. Errol Morris has also made a documentary about the role of photography in the war on terror (*Standard Operating Procedure*).

⁶² Walsh (2011), p. 469.

⁶³ The following observations resonate with, and are inspired by, Timothy Chappell’s insight that ‘the most important thing that distinguishes good ethical thinking about torture from bad ethical thinking about torture is not adherence to any particular moral theory. Rather, the key difference is the deployment of the moral imagination in such thinking. How do we deliberate well about acting badly? A key part of the answer is *with our imaginations engaged*’ (Chappell (2014), p. 43).

⁶⁴ In this article, I am using the philosophical concept of modality in a fairly commonsensical way to contrast hypothetical reasoning with reasoning about actual facts. In epistemology, modality usually refers to what is possible, necessary or impossible. Here, however, I solely focus on how thought experiments construct possible worlds. Consequently, I do not need to commit myself to any specific theory of modality. On the wider debate, see the introduction in Vaidya (2015).

to incorporate in their set-up certain key features that render them recognizable as instantiations of real-world cases from which they imaginatively depart. Thought experiments that are based on what is ‘possible *simpliciter*’ slide into irrelevance precisely because they radically disrupt this crucial link with reality.

Recall that the conceptual framework within which the distinction between this-worldly and other-worldly hypotheticals is made revolves around a democratic orientation of political theory as a civic activity. It follows that drawing the line between this-worldly and other-worldly thought experiments cannot be done in the abstract – it must be undertaken within the context where thought experiments are invoked. We can only establish what is relevant (i.e. possible for us, here and now) by looking at the real-world cases that motivate the creation of thought experiments. There can thus be no other generic yardstick for probing the merits of hypotheticals, apart from their capability of enhancing the judging of real-world problems. It is because imaginary cases ought to be seen through this lens that we need to base our assessment of their virtues and vices on context-specific factors that either facilitate or impede action-guidance.

Framing plays a crucial part in this process. Sontag and Butler have demonstrated that the frames employed to represent violence are themselves the products of decisions about what can, and what cannot, be shown in any given context; what is being left out is as important as what is being revealed. Examining the frames of hypotheticals, their set-up, is so essential because ‘[t]he value of hypotheticals depends on the extent to which they track the critical features of the problem that a moral agent actually faces. To argue from a case that does *not* track the critical moral features of the relevant context disorients both the moral and the legal issues that the hypothetical is designed to illuminate’.⁶⁵ In other words, some frames become unhinged from reality and thereby forfeit their claim to being relevant. Butler’s insistence that precariousness and frailty are distinctive features of human life discloses a criterion by which we may be able to assess the tracking-function of thought experiments.

Exemplars of Irrelevance and Other-Worldliness

Thus far, I have largely refrained from grappling with thought experiments in detail to focus on the theoretical rationale behind their design. But at this stage a closer look at two examples will assist us in more sharply demarcating the boundary between ‘this-worldly’ and ‘other-worldly’ thought experiments. The first is Steinhoff’s discussion of self-defensive rape, which he admits is not for the ‘squeamish’. The reference to rape occurs at a crucial fork in the argument when Steinhoff investigates the permissibility of torture. Recall that the right to self-defence serves as the normative basis for justifying torture: if we grant that it is under certain conditions allowed to kill someone (namely when one needs to defend oneself or an innocent other), then we ought to admit, for exactly the same reason, that torturing someone can be permissible, too.

How about raping somebody, then? Several authors have intimated that pushing the rights-based argument further so as also to include rape would eventually lead to the collapse of the argument from self-defence.⁶⁶ To sustain the coherence of the rights-based defence of torture, Steinhoff is thus forced to claim that the justification from self-defence can be extended to a case like rape. He tries to achieve this by introducing the following thought experiment:

Innocent Jenny, naked in her bedroom, is attacked by Serial Killer, who has broken in. He, too, is naked. Jenny, who is a doctor, is currently treating her vaginal infection with a potent new

⁶⁵ Scheppelle (2005), p. 293.

⁶⁶ See Luban (2014), p. 104; Waldron (2011), pp. 18–19.

ointment, which has the side-effect of killing any man whose penis is exposed to it long and severely enough, something best achieved by sexual intercourse. While the killer is trying to strangle her, they are wrestling on the ground, she gets on top of him, and he gets his hands on her throat and squeezes. In her desperation, she shoves the aggressor's penis – while the aggressor explicitly says 'No!' – into her vagina and starts to move up and down while the man still strangles her. But suddenly the ointment works, the man goes into shock and dies. Jenny is safe.⁶⁷

I shall suppose that it is uncontroversial to deem this an extreme thought experiment. One likely reaction to Innocent Jenny would be to state it is simply too eccentric to tell us anything meaningful about self-defence at all.⁶⁸ The rejoinder could be that philosophers making up hypotheticals like Innocent Jenny are toying around with imaginary cases without acknowledging the severity of the moral problems involved. Innocent Jenny, then, might reveal a disparity between what is morally at stake in a given situation, and how (some) philosophers talk about it. On this account, if someone tried to show, through intricate thought experiments, that the absolute prohibition of torture (or rape) should become an object of discussion, the proper attitude would be to interrupt the conversation and stop deliberating altogether.⁶⁹

Another way of countering a case like Steinhoff's would be to insist that, whilst being peculiar, the thought experiment constructs a situation that is conceivable, but only barely so. The phrase 'only barely so' indicates that the modality underlying Innocent Jenny should be considered problematic. We may hence grant that, in a very distant world, rape could be envisaged as a means of self-defence and that Steinhoff has been successful in imagining such a world. Yet, that statement in itself remains vacuous until we raise a series of subsequent queries: What follows from this hypothetical? What lessons should we draw from it? Where lies its action-guiding element? How, if at all, does it help us make judgements about real-world dilemmas?

The answer to these questions appears to be that, even though Innocent Jenny testifies to Steinhoff's vivid mind-set as regards sexual violence, it does not necessarily succeed in changing established ideas about the wrongness of rape or torture, for that matter. Just because an elaborate counter-example is construable does not mean that a general norm needs revision.⁷⁰ It can be doubted whether many readers of Innocent Jenny would upon reflection come to the conclusion that their pre-conceived intuitions about rape were wrong beforehand – that they should now deem it likely that in situations of self-defence, such as Innocent Jenny, raping somebody would be morally permitted.

To put this point more sharply, as an imaginary case, Innocent Jenny is unhinged from reality and irrelevant precisely because it features an aspect of rape that is entirely absent in all actual cases of rape, namely self-defence. Hence, while it presents us with a scenario that is imaginable and can be said to possess imaginary grip, its practical value for judging real-world cases is negligible. Common sense tells us that in all incidents of real-world rape the characteristic feature of self-defence is absent, which is why Innocent Jenny is not possible for us, here and now. This aspect has ramifications for the debate around torture. If the justification of torture depends, at least partially, on an argumentative analogy between various types of defending

⁶⁷ Steinhoff (2013), p. 149.

⁶⁸ Walsh (2011), pp. 476–7, summarizes this objection effectively. I also take the following reference to Anscombe's work from this paper.

⁶⁹ Elizabeth Anscombe made this point eloquently. See Anscombe (1958), p. 17.

⁷⁰ Timothy Chappell develops an analogous idea when he claims: 'An obvious and constantly recurring type of counter-example to a moral generalization tends to drive a coach and horses through the plausibility of that generalization; a rare and *recherché* type of counter-example need have no such effect' (Chappell (2014), p. 35).

oneself (or innocent others) against aggression, then Innocent Jenny simply does not succeed in establishing a chain of equivalence.

Having grappled with an unambiguously extreme hypothetical, let us now turn to a *prima facie* less outlandish and bizarre one – our second example. It is plausible to presume that the ticking-bomb scenario discussed above is much closer to the world as we know it than Innocent Jenny. Realists like Dershowitz and Gross underscore this idea, and it seems true that the assumptions embedded in it are modally closer to reality than the ones made by Steinhoff. It is also beyond doubt that the ticking-bomb scenario possesses imaginative grip, as various television shows and movies have shown. Should we hence conclude that the ticking-bomb scenario is, using the terminology introduced above, this-worldly and perhaps even benign? Does it articulate a possibility for us, here and now, as Dershowitz and Gross want us to believe?

One way of resisting this conclusion would be to reveal the seriously flawed suppositions on which the ticking-bomb scenario rests. Along these lines, Kim Lane Scheppele demonstrates that the hypothetical fundamentally misconstrues four crucial features of reality that one ought to account for – even in an imaginary world:

- First, the hypothetical assumes that you (as the moral agent to whom the hypothetical is directed) and the terrorist are *alone in the world*. ... There is no institutional context; neither state nor society appears in this picture. But of course in any real-world context, the choice would be made in an institutional setting by those charged with the responsibility to fight terrorism.
- Second, the hypothetical assumes an extraordinary degree of clarity about the situation in which you ... find yourself when the question of whether to torture arises. ... Such certainty may be hypothetically possible, but it will likely never exist.
- Third, the hypothetical assumes that the person to be tortured is *the one* (perhaps even the *only one*) who knows where the ticking bomb is. ... Instead, the more likely question will be whether the person to be tortured really knows anything useful at all.
- Finally, the hypothetical assumes that if the captured person gives you the information after being tortured, the information will in fact be *true* and *useful* in defusing the bomb. Yet torture produces results that are highly unreliable.⁷¹

Should Scheppele's interpretation of the hypothetical's assumptions be accurate, as various authors seem to have confirmed,⁷² we might be inclined to determine that the ticking-bomb scenario shares more with Innocent Jenny than initially suspected. In other words, although it appears to be modally closer to reality, it turns out to be as other-worldly as Steinhoff's hypothetical. On this interpretation, the imaginary case, whilst frequently invoked in the academic as well as the public sphere, is so far detached from what could reasonably be expected to happen in the real world that its merit in guiding action becomes dubious.

The real challenge, then, for those devising imaginary cases is not whether they can conjure up a situation that is within the realm of the possible; with the right kind of training in analytical philosophy it should be expected that they will manage to do so. Rather, for political theorists engaged in the kind of civic activity outlined in the essay's introduction the crux is to what extent the thought experiment helps us navigate the complex moral universe we currently inhabit.⁷³

⁷¹ Scheppele (2005), p. 294.

⁷² See in particular: Brecher (2007); Bufacchi and Arrigo (2006); Shue (2006); Wisniewski (2008).

⁷³ On this point, see the excellent critique in Fried (2014). In his review of Steinhoff's book, Fried makes a creative and helpful distinction between *hypotheticals* and *hypertheticals*. My own suggestion to separate

DIAGNOSING FAILURES AND OUTLINING SUCCESS CONDITIONS

I have intimated that this-worldly hypotheticals are thought experiments that express what is possible for us, here and now; other-worldly hypotheticals, then, are thought experiments that express what is possible *simpliciter*. This division leads to another issue: whether the vices and virtues of hypotheticals depend exclusively on their modal relation to the real world. In other words, are all this-worldly hypotheticals benign and, conversely, are all other-worldly thought experiments pernicious?

The distinction between this-worldly and other-worldly does not fully map on the distinction between benign and pernicious thought experiments. To explain why this is so we must again reflect on action-guidance. As explained above, the capacity of thought experiments to enable judgements is an indicator of their merit. If a hypothetical fails to contribute to our judging of real-world cases due to its irrelevance, it should be considered pernicious.

Nevertheless, crucially, there are two sorts of failure in this regard, which elicit different evaluative responses. An imaginary case might fail to offer action-guidance because it does not succeed in engaging the imagination at all. This is what happens when thought experiments do not manage to develop ‘imaginative grip’ from the outset. A hypothetical might be so outlandish and bizarre that the sheer distance between the possible and the actual world undercuts the readers’ ability to draw any kind of lesson from it. Some extreme hypotheticals resemble science fiction more than anything else, and could therefore be considered entirely benign.⁷⁴ They might not proffer any assistance in judging real-world dilemmas, but neither do they exert any detrimental impact on our ability to judge.

A second sort of failure is more serious for it might destabilize, or even corrupt the ability to judge. Innocent Jenny might be thought to fall into this category. A hypothetical like Steinhoff’s presents us with a case of ‘rape’ that can have a distorting effect on the practical wisdom of those who have learnt to recognize that raping somebody simply cannot be an act of self-defence. While it is, of course, conceivable to envisage ‘rape’ along Steinhoff’s lines, the problem with the thought experiment is not so much that it is purely speculative and possesses no practical value whatsoever when it comes to informing judgements about quandaries in the real world. Rather, one could object that it evokes a possible world in which we are invited to demolish the taboo-like character of rape and, by implication, torture. So, in this case we are dealing with a thought experiment that potentially has ‘imaginative grip’, but its effects on the audience should be deemed deleterious. Innocent Jenny engages the imagination of the readers, but it does so in a problematic fashion.

If neither an extreme hypothetical such as Innocent Jenny nor an apparently less outlandish one like the ticking-bomb scenario manages to sustain a link with the world as we know it, how about this-worldly thought experiments, then? Can there be productive hypotheticals about torture that are relevant to the audience and speak to our world, here and now? As should be obvious from the preceding sections, this article has not opposed the use of thought experiments

(Footnote continued)

analytically ‘possible for us, here and now’ from ‘possible *simpliciter*’ chimes with Fried’s insistence that plausible *hypotheticals* contribute to a concrete and specific interpretation of a shared way of life, while unproductive *hypertheticals* simply stipulate an imaginary scenario that is barely conceivable.

⁷⁴ See, for instance, Thomson’s paper on abortion, where she defends abortion after consensual sex by referring to hypothetical ‘people seeds’, which drift through the air and, although you try to keep them out, still enter into your house through a badly protected window shield. Thompson suggests you are under no obligation to give these people seeds shelter even if you did not take perfect care to protect yourself against their intrusion. By analogy, abortion after consensual sex is morally permissible. See Thomson (1971), p. 59. For an alternative interpretation of Thomson’s thought experiment, see Wisniewski (2008), pp. 111–12.

per se, but rather targeted more specifically those hypotheticals that fail to offer action-guidance. It follows that some imaginary cases, provided they develop imaginative grip, sustain a symbolic link with the real world and enable the judging of difficult situations, must be considered benign. However, to my knowledge such a thought experiment cannot be found in the existing literature. Instead of now attempting to construct a this-worldly hypothetical, I shall outline the success conditions it would have to satisfy to qualify as productive.

The suggestion is that, if one were to transform other-worldly thought experiments about torture into this-worldly ones, one ought to be primarily guided by Butler's admonition to bring 'the human in its frailty and precariousness into view'. As we have observed, Butler stresses the creation and transformation of frames that give prominence to the dignity of victims, no matter who they are or what they have done. Only such novel frames, which often are reactions to hegemonic depictions of muted and invisible suffering, will enable the viewers to express outrage at torture practices and recognize the value of all lives.

This argument is especially poignant when dealing with the clarificationist defence of imaginary cases. Recall Kamm's and McMahan's argument about the advantageous epistemic set-up of hypotheticals, compared with historical or contemporary cases – they strip the real world of superfluous, morally overdetermined information and thus facilitate normative reasoning. The error that many clarificationists commit, however, is that they do not include in their moral universe an acknowledgement of the suffering of those on whom interrogational or self-defensive torture would be exercised. This idea is central for the case argued in this article because it once again underlines the intrinsic connectedness of imaginary cases with the world as we know it. Butler teaches us that frailty and precariousness are features that all human beings share; as such, they can either be accentuated through responsible frames or obfuscated through manipulative ones.

At first sight, this argument seems vulnerable to the objection that clarificationists draw on thought experiments about torture to highlight how much human suffering can potentially be averted by torturing a culpable suspect, be it in self-defence or in defence of innocent others. Hence, a critic could accuse my emphasis on the suffering of those on whom interrogational or self-defensive torture is imposed of misunderstanding the counterfactual nature of the imaginary case.

However, the point to be taken from Butler's discussion is not that the suffering of the tortured must somehow 'trump' the suffering of potential victims; ranking one form of human frailty over another would simply replace a hegemonic hierarchy of grief with an alternative one, which would be a self-defeating enterprise and a mistaken reading of the literature on framing. At no point does this literature intimate we ought to prioritize the fragility of the tortured over the fragility of victims of terrorism. The idea is, rather, to try to set them on an equal footing, which some thought experiments simply do not do. Butler admonishes us that certain hegemonic frames condemn the suffering of some victims to the side-lines, by completely silencing their susceptibility to pain. This is exactly what happens in the vast majority of thought experiments about torture. The very design of these imaginary cases – from the ubiquitous ticking-bomb scenario to Innocent Jenny – makes it impossible to perceive even a trace of the precariousness that characterizes human life.

Obviously, conjuring up hypotheticals that heighten awareness of these aspects of humanity is not an easy exercise. To remain relevant, productive thought experiments about torture would have to attempt to strike a precarious balance between abstracting from the real world, so as to put into clearer focus the complex issue under scrutiny, and maintaining a symbolic link with the world as we know it. The main anchor of this link should be the recognition of human frailty and precariousness, without which the modal distance between real and imagined world

becomes too extensive. Yet, as this section has tried to argue, if we wish to defuse the wholesale scepticism raised by Henry Shue – that artificial cases in general are deleterious to normative reasoning – then we have no choice but to reflect on thought experiments that are possible for us, here and now.

THREE OBJECTIONS AND A CONCLUSION

Three objections can be mounted against this article. The first, concerning its scope, can be dealt with quite swiftly. The perspective from which I have analysed thought experiments is, of course, not unanimously shared by all political theorists. Especially those who subscribe to the view that political theory can and ought to deal in speculation rather than action-guidance will probably find much to disagree with in this essay. This is likely the case due to incompatible starting assumptions. I thus hope that readers adhering to the ‘pure speculation’ strand will be able to appreciate the claims defended in this article against the backdrop of my own fundamental assumptions about the specific role of political theory.

Secondly, clarificationists could counter my proposal by affirming the distinction between theory and practice, as Allhoff does. Their rejoinder might look like this:

Naturally, we believe that torture is terrible and that we would be better off without it; we do not rejoice in the suffering of human beings, after all. But, sadly, we can’t stop here. In order to understand the phenomenon of torture better, we need to know *whether* it is wrong in all circumstances, and that’s where thought experiments can help us a great deal. Hypotheticals free us from the gut reactions we experience when absorbing stories about actually occurring torture practices and thereby allow us to investigate, more rationally and less emotionally, what is morally going on.

There are at least two flaws with this line of explanation. First, given our current knowledge about the pervasiveness of torture practices around the world, maintaining the distinction between moral permissibility and legal authorization, as McMahan for example does, will simply not be enough.⁷⁵ If one is fully aware of the widespread use of torture and if one has learnt to appreciate that torture is abhorrent and condemnable (which in itself must be recognized as significant historical progress), why would one be tempted to make up a modally remote case in which torture were suddenly, under extraordinary circumstances, permissible? What should the audience do with intuitions tested in a situation that shares close to nothing with the real world, such as Innocent Jenny? The determination to separate theory from practice sharply hence sets the debate on the wrong track from the beginning. This article has tried to correct this regrettable mistake.

Secondly, even if we subscribed to the strict distinction between moral knowledge and action, the problems identified above prevent the meaningful exercise of judgement in cases where the modality underpinning the thought experiment is other-worldly. The loss of imaginative grip, the neglect of a concern with human frailty and precariousness, as well as the foreclosing of alternative frames of representation incapacitate the faculty of judging in general, and of judging

⁷⁵ It is a sad sign of the impoverished discourse in some corners of political theory that so little attention is paid to the true horror of torture practices both in democratic and non-democratic societies, whilst endless pages are spent on designing ever more extreme hypotheticals to explore whether torture can be deemed morally permissible. For the best summary work on the complex relationship between torture and democracy, which in many ways could be used as an effective remedy to this overreliance on ideal theory, see the magisterial Rejali (2007).

complex issues of political violence in particular. This is especially clear in a case like Innocent Jenny where the inference of permissible rape crucially depends on a posture of self-defence that is a theoretical possibility, but not a practical one.

One might, thirdly, object to my argument that thought experiments are not only ‘intuition pumps’ – they serve as wake-up calls for reconsidering age-old problems that have become somewhat calcified. The conjurers of what I have named ‘other-worldly’ hypotheticals might hence respond to my criticism that thought experiments, such as Innocent Jenny, are intended to kick-start the debate anew, to make us think harder about the problem of torture. Along Millian lines, they might suggest that the prohibition of torture has become an unexamined dogma, a taboo of sorts, that needs constant probing as much as any other societal norm. Sometimes extreme hypotheticals are indeed defended on the grounds that they estrange readers from the current situation so as to make them perceive more clearly what the values are that inform their practices.⁷⁶ My argument for restraining other-worldly hypotheticals might thus be associated with a conservative spirit, desiring to legislate in favour of certain boundaries of the imagination that ought to remain utterly open.

While such re-imaginings are crucial for societies to explore their core commitments continuously, it matters what readers are being estranged from. If the estrangement reaches a level so radical that the very basis of core commitments becomes shaky, we run the risk of rendering the moral world we inhabit unstable. In theory it is simple to state the paradox this article sought to explore, but in practice hard to resolve it: whereas forcing us to see norms, such as the ban on torture, in a new light is one of the exceptional virtues of such hypotheticals, the drive towards cases that are more and more remote from the world as we know it is one of their inherent vices.

To reiterate a point made earlier: the arguments developed in this article are not directed against all hypotheticals – as I tried to show, the idea that all ‘artificial cases make bad ethics’ is misguided, or at least under-explained. This realization still leaves substantial room for debate around which kinds of imaginary case we wish to promote in political theory. However, I have also suggested that this room is not limitless. Naturally, it remains contestable whether a concrete thought experiment is framed in such a way as to enable judgements of complex cases or not. Yet, an answer to the question whether, for example, Rawls’s ‘original position’ is a productive hypothetical can only be given by examining the concrete context within which the imagination is engaged. Again, the framework undergirding this article does not allow for an *a priori* assessment of all imaginary cases, but it permits us to orient our reflections when we assess the merits of thought experiments in political theory.

This is why it is so vital to uphold the distinction between the two kinds of modality outlined in this article: possible for us, here and now, and possible *simpliciter*. Without it, the imagination is free to flow, but we would be deprived of a means to deliberate on when and why thought experiments go wrong, and, conversely, when and why they help us navigate the moral universe we currently inhabit.

LIST OF REFERENCES

- Alfano, Mark, and Don Loeb. 2014. Experimental Moral Philosophy. In Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*. Available from <http://plato.stanford.edu/archives/sum2014/entries/experimental-moral/>, accessed 20 November 2014.
- Allhoff, Fritz. 2005. A Defense of Torture: Separation of Cases, Ticking Time-Bombs, and Moral Justification. *International Journal of Applied Philosophy* 19:243–64.

⁷⁶ Walsh (2011), pp. 472–3.

- Anscombe, G. E. M. 1958. Modern Moral Philosophy. *Philosophy* 33:1–19.
- Boltanski, Luc. 1999. *Distant Suffering: Morality, Media, and Politics*. Cambridge: Cambridge University Press.
- Bourdieu, Pierre. 2000. *Pascalian Meditations*. Stanford, Calif.: Stanford University Press.
- Brand-Ballard, Jeffrey. 2007. F. M. Kamm: Intricate Ethics: Rights, Responsibilities, and Permissible Harm. *Notre Dame Philosophical Reviews*. Available from <https://ndpr.nd.edu/news/25299-intricate-ethics-rights-responsibilities-and-permissible-harm/>, accessed 14 November 2014.
- Brecher, Robert. 2007. *Torture and the Ticking Bomb*. Malden/Oxford: Blackwell Publishing.
- Brendel, Elke. 2004. Intuition Pumps and the Proper Use of Thought Experiments. *Dialectica* 58:89–108.
- Brown, James Robert, and Fehige, Yiftach. 2016. “Thought Experiments”, *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2016/entries/thought-experiment/>>.
- Bufacchi, Vittorio, and Jean Maria Arrigo. 2006. Torture, Terrorism and the State: A Refutation of the Ticking-Bomb Argument. *Journal of Applied Philosophy* 23:355–73.
- Butler, Judith. 2004. *Precarious Life: The Powers of Mourning and Violence*. London/New York: Verso.
- Butler, Judith. 2007. Torture and the Ethics of Photography. *Environment and Planning D: Society and Space* 25:951–66.
- Butler, Judith. 2009. *Frames of War: When Is Life Grievable?* London/New York: Verso.
- Carroll, Noël. 2002. The Wheel of Virtue: Art, Literature, and Moral Knowledge. *Journal of Aesthetics and Art Criticism* 60:3–26.
- Chambers, Samuel A., and Terrell Carver, eds. 2008. *Judith Butler and Political Theory: Troubling Politics*. London/New York: Routledge.
- Chappell, Timothy. 2014. *Knowing What to Do: Imagination, Virtue, and Platonism in Ethics*. Oxford: Oxford University Press.
- Cohen, G. A. 2008. *Rescuing Justice and Equality*. Cambridge, Mass.: Harvard University Press.
- Cohen, Stanley. 2001. *States of Denial: Knowing about Atrocities and Suffering*. Cambridge/Malden, Mass.: Polity.
- Cova, Florian, and Hichem Naar. 2011. Side-Effect Effect without Side Effects: The Pervasive Impact of Moral Considerations on Judgments of Intentionality. *Philosophical Psychology* 25:837–54.
- Dennett, Daniel C. 1991. *Consciousness Explained*. Boston, Mass.: Little, Brown and Co.
- Dershowitz, Alan M. 2002. *Why Terrorism Works: Understanding the Threat, Responding to the Challenge*. New Haven, Conn.: Yale University Press.
- Edmonds, David. 2013. *Would You Kill the Fat Man? The Trolley Problem and What Your Answer Tells Us about Right and Wrong*. Princeton, N.J.: Princeton University Press.
- Elster, Jakob. 2011. How Outlandish Can Imaginary Cases Be? *Journal of Applied Philosophy* 28:241–58.
- Farrell, Michelle. 2013. The Ticking Bomb Scenario: Origins, Usages and the Contemporary Discourse. Pp. 82–146 in *The Prohibition of Torture in Exceptional Circumstances*. Oxford: Oxford University Press.
- Farrelly, Colin. 2007. Justice in Ideal Theory: A Refutation. *Political Studies* 55:844–64.
- Foot, Philippa. 1967. The Problem of Abortion and the Doctrine of Double Effect. *Oxford Review* 5:5–15.
- Fried, Gregory. 2014. Review of Uwe Steinhoff, *On the Ethics of Torture*, SUNY Press 2013. *Notre Dame Philosophical Reviews*. Available from <http://ndpr.nd.edu/news/48412-on-the-ethics-of-torture/>, accessed 19 May 2014.
- Geuss, Raymond. 2008. *Philosophy and Real Politics*. Princeton, N.J.: Princeton University Press.
- Gilbert, Pablo. 2011. Comparative Assessments of Justice, Political Feasibility, and Ideal Theory. *Ethical Theory and Moral Practice* 15:39–56.
- Gourevitch, Philip. 2008. Exposure: The woman behind the camera at Abu Ghraib. 24 March. Available from <http://www.newyorker.com/magazine/2008/03/24/exposure-5>, accessed 26 November 2014.
- Gross, Oren. 2004. The Prohibition on Torture and the Limits of the Law. Pp. 229–51 in Sanford Levinson, ed., *Torture: A Collection*. Oxford/New York: Oxford University Press.
- Kamm, Frances M. 2006. Terrorism and Several Moral Distinctions. *Legal Theory* 12:19–69.

- Kamm, Frances M. 2008. *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. Oxford/New York: Oxford University Press.
- Knobe, Joshua Michael, and Shaun Nichols, eds. 2008. *Experimental Philosophy*. Oxford/New York: Oxford University Press.
- Laden, Anthony Simon. 2011. The Key to/of Public Philosophy. *Political Theory* 39:112–17.
- Lebow, Richard Ned. 2010. *Forbidden Fruit: Counterfactuals and International Relations*. Princeton, N.J.: Princeton University Press.
- Luban, David. 2009. Human Dignity, Humiliation, and Torture. *Kennedy Institute of Ethics Journal* 19:211–30.
- Luban, David. 2014. *Torture, Power, and Law*. Cambridge/New York: Cambridge University Press.
- McMahan, Jeff. 2002. *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press.
- McMahan, Jeff. 2008. Torture in Principle and in Practice. *Public Affairs Quarterly* 22:91–108.
- McMahan, Jeff. 2009. *Killing in War*. Oxford: Clarendon Press.
- McMahan, Jeff. Forthcoming. Torture and Method in Moral Philosophy. In *Torture and the Rule of Law*, edited by Scott Anderson and Martha Nussbaum. Chicago: Chicago University Press.
- Mills, Charles W. 2005. 'Ideal Theory' as Ideology. *Hypatia* 20:165–83.
- Miščević, Nenad. 2013. Political Thought Experiments from Plato to Rawls. Pp. 191–206 in Melanie Frappier, Letitia Meynell and Robert Brown, eds, *Thought Experiments in Science, Philosophy, and the Arts*. New York: Routledge.
- Morris, Errol. 2008. The Most Curious Thing. *New York Times* blog. Available from <http://morris.blogs.nytimes.com/2008/05/19/the-most-curious-thing/>, accessed 26 November 2014.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- Rawls, John. 1999. *A Theory of Justice*, Revised Edition. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Rejali, Darius M. 2007. *Torture and Democracy*. Princeton, N.J.: Princeton University Press.
- Scheppele, Kim Lane. 2005. Hypothetical Torture in the War on Terrorism. *Journal of National Security Law & Policy* 1:285–340.
- Shapiro, Ian. 2005. *The Flight from Reality in the Human Sciences*. Princeton, N.J.: Princeton University Press.
- Shue, Henry. 1978. Torture. *Philosophy and Public Affairs* 7:124–43.
- Shue, Henry. 2006. Torture in Dreamland: Disposing of the Ticking Bomb. *Case Western Reserve Journal of International Law* 37:231–9.
- Sontag, Susan. 1977. *On Photography*. New York: Farrar, Straus and Giroux.
- Sontag, Susan. 2003. *Regarding the Pain of Others*. New York: Farrar, Straus and Giroux.
- Sontag, Susan. 2004. Regarding the Torture of Others. *New York Times*, 23 May. Available from <http://www.nytimes.com/2004/05/23/magazine/regarding-the-torture-of-others.html>, accessed 9 October 2014.
- Sorensen, Roy A. 1992. *Thought Experiments*. New York/Oxford: Oxford University Press.
- Standard Operating Procedure*, directed by Errol Morris (2008), available at <http://www.imdb.com/title/tt0896866/>
- Steinhoff, Uwe. 2010. Defusing the Ticking Social Bomb Argument: The Right to Self-Defensive Torture. *Global Dialogue* 12:1 Winter/Spring. Available from <http://www.worlddialogue.org/content.php?id=457>, accessed 19 November 2014.
- Steinhoff, Uwe. 2013. *On the Ethics of Torture*. Albany: State University of New York Press.
- Stone, Alison. 2005. Towards a Genealogical Feminism: A Reading of Judith Butler's Political Thought. *Contemporary Political Theory* 4:4–24.
- Tetlock, Philip E., and Aaron Belkin, eds. 1996. *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*. Princeton, N.J.: Princeton University Press.
- Thomson, Judith Jarvis. 1971. A Defense of Abortion. *Philosophy and Public Affairs* 1:47–66.
- Tully, James. 2008. *Public Philosophy in a New Key*. Cambridge: Cambridge University Press.

- Vaidya, Anand. 2015. The Epistemology of Modality. In Edward N. Zalta ed., *The Stanford Encyclopedia of Philosophy*. Available from <http://plato.stanford.edu/archives/sum2015/entries/modality-epistemology/>, accessed 6 October 2015.
- Valentini, Laura. 2012. Ideal vs. Non-Ideal Theory: A Conceptual Map. *Philosophy Compass* 7:654–64.
- Waldron, Jeremy. 2011. *What Are Moral Absolutes Like?* Rochester, N.Y.: Social Science Research Network. Available from <http://papers.ssrn.com/abstract=1906850>, accessed 26 September 2013.
- Walsh, Adrian. 2011. A Moderate Defence of the Use of Thought Experiments in Applied Ethics. *Ethical Theory and Moral Practice* 14:467–81.
- Walzer, Michael. 2006. *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, 4th edn. New York: Basic Books.
- Williams, Bernard. 2005. *In the Beginning Was the Deed: Realism and Moralism in Political Argument*. Princeton, N.J.: Princeton University Press.
- Wisnewski, J. Jeremy. 2008. It's About Time: Defusing the Ticking Bomb Argument. *International Journal of Applied Philosophy* 22:103–16.
- Wisnewski, J. Jeremy. 2014. Review of F. M. Kamm, *Ethics for Enemies: Terror, Torture, and War*, Oxford University Press, 2011, in *Journal of Moral Philosophy* 11:657–60.
- Zehfuss, Maja. 2009. Hierarchies of Grief and the Possibility of War: Remembering UK Fatalities in Iraq. *Millennium: Journal of International Studies* 38:419–40.