Chapter 5 :   Ticking Time-Bomb Methodology

In thinking about putative justifications for interrogational torture, we are commonly asked to imagine exceptional cases wherein such torture is necessary to save some significant number of lives. These cases are generally referred to as ticking time-bomb cases, which invites us to think of the relationship they bear to terrorism: some terrorist has planted a bomb in a crowded metropolitan center which will kill many noncombatants unless the terrorist is tortured. But, as I suggested in the first paragraph of the last chapter, terrorism need not have anything to do with these cases nor, really, do bombs. Rather, what matters is that there is some threat to many people which can only be avoided— and which, in most formulations, certainly will be—through the torture of someone already in custody. The "ticking time-bomb" locution is therefore somewhat narrow, but not in any drastically misleading way. Furthermore, most of the contexts worth considering—by which I mean real-world ones most closely approximating these hypothetical constructs—probably will be those involving terrorists and weapons, if not necessarily bombs. And, regardless, there is merit in following the standard usage. Having thus registered these disclaimers, I shall hereafter revert to that standard usage.

The purpose of this chapter is methodological rather than normative: by this distinction, I mean that we will herein consider the logic and deployment of ticking time-bomb cases rather than what follows from them (i.e., whether they justify interrogational torture in exceptional circumstances). My contention is that, despite their simplicity, these cases have been misunderstood, whether in what they claim, what role they are should to play in our moral thinking, or even whether should play any role at all. The next chapter, by contrast, will be normative: having thought about how to properly understand ticking time-bomb cases, we shall then use them to derive normative conclusions. But it is important to

first get clear on the methodology so that these subsequent discussions can be situated on the requisite foundations.

5.1   Origins of the Ticking Time-Bomb

Let us start by considering the origins of ticking time-bomb cases; in doing so, we will also start to see some of the different formulations that these cases assume. In terms of the philosophical literature, an early formulation owes to a seminal essay by Henry Shue:

> [S]uppose a fanatic, perfectly willing to die rather than collaborate in the thwarting of his own scheme, has set a hidden nuclear device to explode in the heart of Paris. There is no time to evacuate the innocent people or even the movable art treasures—the only hope of preventing tragedy is to torture the perpetrator, find the device, and deactivate it.[1]

But not only have these cases appeared in academic journals, they have also crossed over to popular media outlets and, thereafter, into public consciousness; this is noteworthy. For example, consider the following, which comes from an essay Michael Levin wrote in *Newsweek*:

> Suppose a terrorist has hidden an atomic bomb on Manhattan Island which will detonate at noon… Suppose, further, that he is caught at 10 a.m…., but preferring death to failure, won't disclose where the bomb is. What do we do? If we follow due process, wait for his lawyer, arraign him, millions of people will die. If the only way to save those lives is to subject the terrorist to the most excruciating possible pain, what grounds can there be for not doing so? I suggest that there are none.[2]

While less-often acknowledged, Anthony Quinton had, over a decade earlier, briefly entertained a similar scenario in *The Listener*, a now-defunct weekly magazine established by the BBC:

> I do not see on what basis anyone could argue that the prohibition of torture is an absolute moral principle…Consider a man caught planting a bomb in a large hospital, which no one but he knows how to defuse and no one dare touch for fear of setting it off. It was this kind of extreme situation that I had in mind when I said earlier that I thought torture could be justifiable.[3]

---

[1] Henry Shue, "Torture," *Philosophy and Public Affairs* 7.2 (1978), p. 141 [124-143].
[2] Michael Levin, "The Case for Torture," *Newsweek* (February 7, 1982): 7. Levin is a philosopher, albeit one of the few to be published in *Newsweek*.
[3] Anthony Quinton, "Views," *The Listener* (December 2, 1971): 757-758. I thank Bob Brecher for directing me to this source and Ryan Pflum for being able to track it down.

And, before that, Jean Lartéguy fictionalized an episode from the French occupation of Algeria in his 1960 book, *Les Centurions*.[4] But, despite these formulations of ticking time-bomb cases in the second half of the 20th century, Jeremy Bentham had been writing on the morality of torture almost two hundred years earlier, dating from the late 1770s.[5] (He also provided one of the earlier characterizations of torture.[6]) Bentham probably had the first formulation of a case that looked anything like a ticking time-bomb case, though this came later—in 1804—and was not otherwise attached to a systematic treatment of torture. Consider what he wrote:

> Suppose an occasion, to arise, in which a suspicion is entertained, as strong as that which would be received as a sufficient ground for arrest and commitment as for felony—a suspicion that at this very time a considerable number of individuals are actually suffering, by illegal violence inflictions equal in intensity to those which if inflicted by the hand of justice, would universally be spoken of under the name of torture. For the purpose of rescuing from torture these hundred innocents, should any scruple be made of applying equal or superior torture, to extract the requisite information from the mouth of one criminal, who having it in his power to make known the place where at this time the enormity was practicing or about to be practiced, should refuse to do so? To say nothing of wisdom, could any pretense be made so much as to the praise of blind and vulgar humanity, by the man who to save one criminal, should determine to abandon [one hundred] innocent persons to the same fate?[7]

The principal difference between Bentham's case and the others previously presented is simply whether the harm that the torture aims to dispel is already active (*viz.*, the current torture of innocents) or else prospective (*viz.*, the future explosion of a bomb). Morally, there need not be any difference between

---

[4] Jean Lartéguy, *Les Centurions*, trans. Xan Fielding (New York: Avon Books [1960] 1961). For discussion, see Darius Rejali, *Torture and Democracy* (Princeton, NJ: Princeton University Press, 2007), pp. 535, 545-548, 775 n. 51.

[5] Jeremy Bentham, "Of Torture," Bentham MSS Box 46 (1786-1780?), 63-70, 56-62. [Note that all Bentham citations are to "folios" rather than page numbers; the p./pp. should therefore be omitted.] Reprinted in W.L. Twining and P.E. Twining, "Bentham on Torture," *Northern Ireland Legal Quarterly* 24 (1973): 307-356. All citations will be to the Twining and Twining essay which is more widely available than the archives. These principal writings on torture contain two fragments, though Twining and Twining suggest that the second fragment was actually written first; I agree that it is more useful to read the pages in the order suggested (63-70, 56-62) rather than the alternative (e.g., 56-70). Other fragments also exist wherein Bentham discusses torture; see MSS Box 74.b (1804), 405-406, 414, and 426-429. See also MSS Box 99 (1787-1789?), 173. Bentham's published work barely mentions torture, aside from a few casual references. See Jeremy Bentham, *The Works of Jeremy Bentham*, 11 vols., ed. John Bowring (Edinburgh: William Tait, 1838-1843), I.231, 393, 414, 424-444; IV.211; VII.454-455, 522-523, 525. All references adapted from Twining and Twining (1973), p. 307.

[6] "Torture, as I understand it, is where a person is made to suffer any violent pain of body in order to compel him to do something or to desist from doing something which done or desisted from the penal application is immediately made to cease" (Twining and Twining 1973, p. 309).

[7] Jeremy Bentham, Bentham MSS Box 74.b, 429 (May 27, 1804). Quoted in Twining and Twining (1973), p. 347.

these cases: what matters is whether the torture is necessary to prevent the harm. If that harm is temporally distant, then that would undermine the need to torture insofar as there might be other—and less morally offensive—ways to dispel it. But so long as the torture is necessary, then whether the harm is ongoing, imminent, or even temporally distant is irrelevant, at least so long as 'necessary' really means what it says. We will return to this below but, for now, the point is merely that Bentham's case is structurally similar to the others.

But aside from this more casual presentation of a single case, Bentham also offered a more extended treatment of torture, as recorded in two manuscript fragments. It is worth considering these fragments for at least three reasons, only one of which is historical. The second is more philosophical in that Bentham starts to elucidate some of the key logical elements of ticking time-bomb thinking, even if that discussion floats free of a particular ticking time-bomb case. And the third bears on the relationship between ticking time-bomb methodology and utilitarianism, a relationship that is more complicated than is usually acknowledged. With these three reasons in mind, let us now look at some of Bentham's writings on torture.

First, Bentham asserts that torture may be applied in two cases. "The first is where the thing which a Man is required to do being a thing which the public has an interest in his doing, is a thing which for a certainty is in his power to do."[8] And he continues that torture is otherwise permissible:

> …where a man is required what probably though not certainly is in his power to do; and for the not doing of which it is possible that he may suffer, although he be innocent; but which the public has so great an interest in his doing that the danger of what may ensue from his not doing it is a greater danger than even that of an innocent person's suffering the greatest degree of pain that can be suffered by Torture, of the kind and in the quantity permitted to be employed.[9]

Then Bentham asks: "Are there in practice any cases that can be ranked under this head? If there be any, it is plain that there can be very few."[10] That Bentham was reserved about the extent to which

---

[8] Twining and Twining (1973), p. 312.
[9] Twining and Twining (1973), p. 312-313.
[10] Twining and Twining (1973), p. 313.

torture can be justified is noteworthy: being a utilitarian hardly commits one to the promiscuous use of torture as there are myriad utilitarian reasons to oppose it.[11] We will discuss some of these below and throughout the rest of the book, but I want to get early purchase on the concept of exceptional—as opposed to normalized—torture.[12]

After these introductory remarks, Bentham goes on to offer a series of moral rules that have to be satisfied for the legitimate application of torture. While the details of those rules need not concern us here, suffice it to say that they are precisely the sorts of principles that undergird contemporary ticking time-bomb cases. For example, torture should not be applied without (near-) certainty that the would-be tortured has the relevant knowledge (Rule 1); that torture is only appropriate as a last resort in "cases which admit of no delay" (Rule 3);[13] that minimal means should always be preferred to extreme ones (Rule 4); that the prospective benefits are greater than the prospective costs (Rules 5 and 7), and so on.[14]

Importantly, many of these rules were effectively codified in Bentham's hedonic calculus, published shortly thereafter.[15] Bentham predicated his utilitarianism on seven factors, all of which are at least implicitly manifest in the ticking time-bomb cases: intensity, duration, certainty (or uncertainty), propinquity (or remoteness), fecundity, purity, and extent. While intensity and duration are rarely

---

[11] See also Twining and Twining (1973), p. 348-350.

[12] Some critics allege that this distinction is untenable and that exceptional torture would inevitably result in normalized torture. We will return to this issue in §§7.2-7.3.

[13] As I read them, Rule 3 is curiously at odds with the forthcoming Rule 5. First, consider Rule 3, which says that torture "ought not to be employed but in cases which admit of no delay." And now, Rule 5: "Even on occasions which admit not of delay, [torture] ought not be employed but in Cases where the benefit produced by doing of the thing required is such as can warrant the employing of so extreme a remedy." Twining and Twining (1973), p. 313. The straightforward part of Rule 5 is that torture should not be practiced when the costs outweigh the benefits. But even if the benefits outweigh the costs, it does not follow that torture should be practiced since there might be some other way to achieve a better net outcome. In cases which admit of delay, one might as well wait before torturing since this alternative might present itself; it is only when the cases do not admit of delay that such alternatives can be ruled out. It seems to me that Bentham had it right with Rule 3 and should not have thereafter gone on to contradict—or at least weaken—it in Rule 5.

[14] Twining and Twining (1973), p. 312-15.

[15] Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (Mineola, NY: Dover Publications, Inc., [1780] 2007), esp. ch. 4.

emphasized in the cases, they certainly *could* be, and such invocations would seemingly only make the cases more efficacious: imagine that the terrorist need only be subject to a "comparatively minor and brief" form of torture to disclose the location of the bomb.[16] (Note that this is not to suggest that torture could ever be minor—which some might argue to be incoherent—but rather that it most certainly comes in degrees and could be *comparatively* minor.)

The other features, though, I regard as at least near-explicit in the cases. Certainty is perhaps the most conspicuous feature of ticking time-bomb cases: *everything* is certain. It is certain that the detainee is a terrorist. It is certain that he has information regarding the location of the bomb. It is certain that the torture will produce the information. It is certain that the information will lead to the timely deactiviation of the bomb. And many critics of the cases promptly seize upon all of this certainty, which undoubtedly represents a departure from (at least almost all) actual cases;[17] I will return to these criticisms below and in following chapters.

Next come fecundity and purity, which are opposite sides of same coin: when we torture, we will get good things (fecundity) and we will not get bad things (purity). This fecundity is thereafter magnified by the invocation of extent, the last of Bentham's elements, which holds that, not only will a single life be saved through the torture, but rather *a lot* of lives will be saved (cf., Levin's "millions"). The purity condition comes through insofar as no bad consequences—aside from the pain and suffering of the tortured (see §4.3)—are postulated. And, while it is open for the critic to say that such cases do not preclude such consequences, it is equally open to the proponent to merely issue such a stipulation, at least at this stage of the dialectic.

---

[16] Note that Levin goes the other way and even assumes that the terrorist be subjected to the "most excruciating possible pain" (Levin 1982).
[17] See, for example, Henry Shue, "Torture in Dreamland: Disposing of the Ticking Bomb," *Case Western Reserve Journal of International Law* 37.**2**&3 (2006): 231-239.

Critics nevertheless do complain about ticking time-bomb cases precisely on the issue of purity. They assert: that torture would have to be institutionalized,[18] including the implementation of training programs for the torturers;[19] that such institutionalization portends harms for liberal democracies;[20] that us torturing our enemies makes it more likely that our enemies will torture us;[21] that torture makes it more likely for us to perpetuate other wrongs,[22] and so on.[23] We will return to these issues in the next chapter but, for now, I just want to mention some of them.

Regardless, the proponent of ticking time-bomb methodology is still free to say "look, that just *is not how the case goes*!"; the critic cannot load conditions into the case that are patently excluded by presupposition.[24] To do so is simply to change the case and to ask a different question altogether, and precisely not the one that we currently care about. Rather, the question at hand is whether torture is permissible given features either stated or implied in the ticking time-bomb cases, and this is a question on which moral philosophy owes an answer. Following that inquiry (Chapters 6), we can then think

---

[18] Michael Davis, "The Moral Justifiability of Torture and other Cruel, Inhuman, or Degrading Treatment," *International Journal of Applied Philosophy* 19.2 (2005): 161-178.

[19] Jessica Wolfendale, "Training Torturers: A Critique of the "Ticking-Bomb Argument," *Social Theory and Practice* 32.2 (2006): 269-287.

[20] David Luban, "Liberalism, Torture, and the Ticking Bomb," *Virginia Law Review* 91.6 (2005): 1425-1461. See also Vittorio Bufacchi and Jean Maria Arrigo, "Torture, Terrorism and the State: A Refutation of the Ticking Time-Bomb Argument," *Journal of Applied Philosophy* 23.3 (2006): 355-373.

[21] Jean Maria Arrigo, "A Utilitarian Argument against Torture," *Science and Engineering Ethics* 10.3 (2004): 1-30. See also Matthew Wynia, "Consequentialism and Harsh Interrogations," *American Journal of Bioethics* 5.1 (2005): 4-6.

[22] Andrew Fiala, "A Critique of Exceptions: Torture, Terrorism, and the Lesser Evil Argument," *International Journal of Applied Philosophy* 20.1 (2006): 127-142. See also John Kleinig, "Ticking Bombs and Torture Warrants", *Deakin Law Review* 10.2 (2006): 614-627.

[23] For a thorough discussion on many of these issues, see Bob Brecher, *Torture and the Ticking Bomb* (Malden, MA: Blackwell Publishing, 2007), esp. chaps. 2-3.

[24] Note that there are some analogies with the critics' responses here and those elicited through Jonathan Haidt's work: even when some morally salient features are ruled out *ex ante*, subjects nevertheless try to rationalize their preferred conclusion *ex post* by searching for a plausible justification.

In one of Haidt's cases, subjects are asked to consider a brother and sister who have had sex, and then to explain whether and why they find the act wrong. The details of the case are such to preclude any of the answers given: the siblings were careful about contraception; they keep it a secret; they enjoy it and it makes them closer, and so on. Nevertheless, subjects continue to express disapprobation and, after being pointed out that none of their reasons is applicable, often revert to something like "still, it's just wrong." See Jonathan Haidt, "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108 (2001): 814-834.

about what implications it has vis-à-vis (real world) cases that relax some of the idealizations and abstractions (Chapters 7). But, as the dialectic goes, the aforementioned complaints are completely irrelevant.

To return to a distinction from the introduction to this chapter, the present aims are methodological rather than normative: I do not purport to have even asked questions about the morality of torture in the real world, but rather to be defending a certain methodological approach to those questions. This approach is one in which ticking time-bomb cases figure prominently, regardless of the relationship that they bear to real-world cases. The next chapter will be normative, though will stay within the realm of this philosophical construct; the third part of the book will then move to the real world. But, for now, let us first get clear about how ticking time-bomb methodology is meant to work lest we not properly understand its application. For the remainder of this chapter, I propose to first consider objections against this proposal, as well as to further articulate the logic and underpinnings of ticking time-bomb methodology.

5.2   Intuitions and Thought Experiments

Having seen some of the origins of the ticking time-bomb cases, as well as their putative grounding in classical utilitarian thinking, we can turn to the role these cases are meant to play in the torture debate. Despite the frequency with which the cases are invoked, I think that their purpose has often been misunderstood. The most obvious point is that the cases are constructed such as to elicit intuitions about the moral permissibility of torture in rarefied situations. Or, to put it another way, they are meant to propose sufficient conditions for torture's permissibility: *if* we had a scenario in which ticking time-bomb conditions met, then torture is permissible.

For this reason, it hardly follows that ticking time-bomb cases provide too stringent of requirements for the application of torture—i.e., that they "[set] the bar too high"[25]—or that torture will never be allowed in the real world because these stipulated conditions will never be met. These interpretations commit the fallacy of denying the antecedent: given the conditional that "if ticking time-bomb conditions are met, then torture is morally permissible", nothing follows about the permissibility of torture when those conditions are not met in just the same way that "if today is Tuesday, then I am wearing a green shirt" says nothing about the color of my clothing on days other than Tuesday. To say this another way, ticking time-bomb cases must be providing sufficient—as opposed to necessary—conditions for the permissibility of torture since, if the conditions were (merely) necessary, then that permissibility would never follow insofar as there might be additional necessary conditions beyond those specified.[26] Clearly ticking time-bomb cases invite us to think that torture is morally permissible under the appropriate conditions, so the logic of these cases has to be one of sufficiency rather than necessity.

Unlike some other authors, I think in terms of ticking time-bomb *cases* as opposed to ticking time-bomb *arguments*.[27] Arguments have premises and conclusions, and ticking time-bomb cases do not. Rather, they are thought experiments which are meant to elicit our intuitions in regards to some proposed scenario. They tend to be framed in terms of things we should "imagine" or "suppose", rather than in terms of premises. The difference between these two is that premises have truth values, whereas "imagine that P" or "suppose that Q" do not (i.e., these are imperatives, like "close the door"). Looking at the formulations by Shue, Levin, and Quinton above, I read them all as cases, though Levin

---

[25] Kleinig (2006), p. 614.

[26] For example, consider that someone has to be at least 35 years old to be eligible as President of the US. But satisfying this condition is insufficient for eligibility since there are two additional requirements: being a natural-born citizen and having lived in the US for fourteen years. Therefore, we cannot infer anything about anyone's eligibility merely on the grounds that s/he has satisfied the age requirement. Analogously, if ticking time-bomb conditions were necessary for the permissibility of torture, we still would not know whether torture was permissible, since there could be other conditions required for (joint) sufficiency, as in the case of those for presidential eligibility.

[27] Cf., e.g., Bufacchi and Arrigo (2006).

and Quinton go on to say what they think about the permissibility of torture given the relevant suppositions. And this is as it should be: the cases invite moral reflection.

Importantly, ticking time-bomb methodology does not encourage us to infer the permissibility of torture from our intuitions alone. "I intuit P, therefore P" is a bad argument because our intuitions are (morally) fallible; or, to put it another way, we should not read off normative conclusions from descriptive premises (*viz.*, ones about our intuitions). Rather, ticking time-bomb cases reveal our deepest moral commitments—in the specified cases only—and those commitments then figure into our moral theorizing. Sometimes those commitments can be sustained through this process, but sometimes they are rejected given conflicting moral principles, other moral commitments, and so on.[28] This is the normative—as opposed to methodological—project that will occupy us in the next chapter. But, for now, I propose that that ticking time-bomb cases are meant to offer conditions under which we are invited to countenance the moral permissibility of torture.

When critics have properly understood the logic of ticking time-bomb methodology, they have taken a range of responses, ranging from dubious to downright hostile (cf., David Luban's allegations of "intellectual fraud"[29]). Still, very few authors actually engage the cases on their own terms, instead offering empirical critiques of the presuppositions. These will have to be dealt with in later chapters but, again, the purpose of the current chapter is wholly methodological. One of the few purely methodological critiques is offered in an ambitious paper by Michael Davis, whose ideas we also met in §4.3. For present purposes, one reason that Davis's paper is so useful is that he analyzes ticking time-bomb methodology in the same way that I do, though we reach different conclusions: I think that this methodology is useful, whereas he thinks it should be discarded. In particular, we both agree that the cases are meant to elicit intuitions, as opposed, for example, to thinking that they are meant to

---

[28] I have in mind something like John Rawls's reflective equilibrium, though the particular details—whether Rawlsian or otherwise—are not important for present purposes. See John Rawls, *A Theory of Justice*, rev. ed. (Cambridge, MA: Belnap Press, 1999), pp. 18-19, 42-45.

[29] Luban (2005), p. 1427.

constitute arguments. Davis then has three objections to the intuitions that the cases supposedly elicit,

all of which threaten their role in some broader sort of moral theorizing. He summarizes these from the

outset, then expands upon them later:

> The intuition [that torture is permissible in those cases] is not as widely shared as necessary to constitute the required demonstration. Second, the intuition is not as reliable as necessary for such a demonstration. We lack the experience that would vouch for it. And, third…what we are intuiting…is an excuse rather than a justification.[30]

Suffice it to say, I think that all of these objections are misguided. Regarding the claim that the intuition

is not as widely shared as necessary, there is simply no empirical data to back up his claim; he merely

offers passing allusions to "[g]ood Kantians" and "[s]ome Catholics."[31] In §5.3, I will present some

empirical data about intuitions in ticking time-bomb cases, and I take that data to flatly controvert

Davis's claim. But my presentation of that data is not really even meant to have normative implications,

precisely for the reason mentioned earlier: the fact that some people intuit P does not necessarily mean

that P is true. Perhaps out of charity, Davis seems to consider the idea that a widely-shared intuition

would have justificatory power, though I reject this idea. And, furthermore, he seems to think that if

that intuition is not universal, it cannot have justificatory power. Even if I did think that intuitions had

the justificatory potential that he entertains, I would reject this requirement: the fact that racists, for

example, have certain moral intuitions does not mean that the intuitions of the rest of us are any less

likely to track moral truth, even if intuition alone is insufficient to identify that truth. Regardless, I reject

the idea that widely-shared intuition is necessary for moral justification. (As a philosopher whose views

bear little intuitive popularity, I had certainly better.)

The second objection, though, really threatens: it alleges that the intuitions we have in ticking

time-bomb cases are irrelevant because these cases fall so far from our experience. This objection is

importantly different from those of other critics who simply deny the empirical plausibility of ticking

---

[30] Davis (2005), p. 161.
[31] Davis (2005), p. 171.

time-bomb suppositions since those objections allow for the legitimacy of ticking time-bomb intuitions, yet deny their relevance. Moves are available to the proponent of ticking time-bomb methodology here, such as to argue for that relevance. Davis's objection, though, cuts deeper by threatening to take those intuitions off the table altogether. My reply to this takes two parts: first, I question whether his claim is true and then, granting it, I argue that it is problematic.

Is it true that we lack whatever experience could support ticking time-bomb intuitions? For the sake of argument, let us grant that none of us has had the experience of torturing a terrorist in order to prevent a threat; maybe some of the readers have been less fortunate, but that need not concern us here. Nevertheless, I submit that we do not need experience of ticking time-bomb cases in order to have whatever experience is relevant. Consider, for example, the following thought experiment:

> *Lost Wallet*. One Tuesday, you are walking home and see a brown wallet sitting on the sidewalk. You pick it up and find $240 inside, as well as the driver's license of its owner. It turns out that the owner lives two houses down, and he happens to be sitting on his front porch. He seems well-enough off, though, and you have just lost your job, thus threatening the well-being of your family. The $240 could be used to buy groceries for a week, as well as to purchase important medicine for one of your ailing children, none of which you can otherwise afford.

Should you return the wallet? For this discussion, the answer is irrelevant; the point is to question whether we have the necessary experience to legitimize our intuitive responses to the case, whatever they are. Strictly speaking, no: this exact situation has almost certainly never happened to any of us. Does that mean that the informativeness of our intuitions is somehow compromised? Absolutely not. Those intuitions are still morally informative, despite the fact that they are not narrowly grounded in experience. Consider that some of us have found wallets, but it was Wednesday, the wallet was black, or whatever. Those experiences share the *relevant* moral features insofar as days of the week, color, and so on are morally irrelevant. But what if some of us have never found a wallet at all? Or had ailing children? For example, I do not even have children at all, yet it certainly seems like hunger and sickness are things that I understand. The morally relevant features of this case are things like property rights, honesty, suffering, and so on; however they are instantiated is unimportant. All of us have experience

with those things, so therefore our intuitions are useful. It is actually hard for me to think of any thought experiment that postulates morally relevant features with which none of us could identify, but this difficulty is precisely to make the point: we all have rich moral lives and experiences that comprise myriad moral features. If we did not, there would something oddly obscure about those proposed moral features.

I take all of the above to be straightforward so, for the objection to have any plausibility, it would have to hold that ticking time-bomb cases are so "radically different" from our experience as to significantly distinguish them from the lost wallet case; the lost wallet case would have to be pedestrian or quotidian by comparison. But how would this line be motivated? In a straightforward sense, the cases bear the same moral features. The owner of the wallet has a right to have it back and, let us suppose, the terrorist has a right not to be tortured. By violating that right, some greater suffering is alleviated. So what is the difference? It could be pragmatic insofar as ticking time-bomb cases portend some sort of practical abuse, but I find that to be irrelevant to philosophical methodology. Is it that there is a *lot* of suffering (cf., thousands of lives) and all these lives are somehow distortional (i.e., that the number of lives somehow "messes up" our intuitions)? This cannot be right for the simple reason that the number of lives *matters*: torture is morally better (or less morally bad) if it saves thousands of people than if it saves two people.[32] At any rate, I doubt that ticking time-bomb cases are radically different from those in our experiential repertoire or, if the difference is only a matter of degree, that degree is insufficient to invalidate the intuitions.

But suppose that you remain unconvinced. Consider some of the greatest thought experiments in moral philosophy. Start with Judith Jarvis Thomson's famous violinist case, which is so fantastic that it is worth quoting at length:

---

[32] Or at least I take this to be self-evident. For a controversial dissent, see John Taurek, "Should the Numbers Count?" *Philosophy & Public Affairs* 6 (1977): 293-316. But even if Taurek thinks that we should flip a coin to see whether to choose between saving one person or else five others, he could still agree with my claim.

> [L]et me ask you to imagine this. You wake up in the morning and find yourself back to back in bed with an unconscious violinist. A famous unconscious violinist. He has been found to have a fatal kidney ailment, and the Society of Music Lovers has canvassed all the available medical records and found that you alone have the right blood type to help. They have therefore kidnapped you, and last night the violinist's circulatory system was plugged into yours, so that your kidneys can be used to extract poisons from his blood as well as your own. The director of the hospital now tells you, "Look, we're sorry the Society of Music Lovers did this to you—we would never have permitted it if we had known. But still, they did it, and the violinist is now plugged into you. To unplug you would be to kill him. But never mind, it's only for nine months. By then he will have recovered from his ailment, and can safely be unplugged from you." Is it morally incumbent on you to accede to this situation?[33]

Thomson's case is meant to be analogous to pregnancy due to rape, and our intuition in this case that it is permissible to disconnect the violinist is meant to similarly suggest that it is morally permissible to terminate rape-induced pregnancies.

Or else take Peter Singer's shallow pond case, which asks us to imagine whether we are obligated to save a child drowning in a shallow pond even if it would be inconvenient and lead to the muddying of our clothes.[34] Singer proposes this case as analogous to that of our obligations vis-à-vis easily preventable deaths of the world's poor and takes our intuitions in the shallow pond case to inform those obligations. Or, finally, consider Philippa Foot's trolley case—used to motivate the doctrine of double effect—retold to greater effect by Thomson:

> Suppose you are the driver of a trolley. The trolley rounds a bend, and there come into view ahead five track workmen, who have been repairing the track. The track goes through a bit of a valley at that point, and the sides are steep, so you must stop the trolley if you are to avoid running the five men down. You step on the brakes, but alas they don't work. Now suddenly you see a spur of track leading off to the right. You can turn the trolley onto it, and thus save the five men on the straight track ahead. Unfortunately…there is one track workman on the spur of the track. He can no more get off the track in time than the five can, so you will kill him if you turn the trolley onto him. Is it morally permissible for you to turn the trolley?[35]

---

[33] Judith Jarvis Thomson, "A Defense of Abortion," *Philosophy & Public Affairs* 1.1 (1971), pp. 48-49 [47-66].

[34] Peter Singer, "Famine, Affluence, and Morality," *Philosophy & Public Affairs* 1.3 (1972), p. 231 [229-243].

[35] Judith Jarvis Thomson, "The Trolley Problem," *Yale Law Journal* 94.6 (1985), p. 1395 [1395-1415]. See also Judith Jarvis Thomson, "Killing, Letting Die, and the Trolley Problem," *Monist* 59 (1976): 204-217. Thomson even acknowledges that "most of [her] examples are just long-winded expansion's of [Foot's]." Thomson (1976), p. 217 n. 2. While the trolley problem is attributed to Foot, her presentation is actually somewhat different: imagine "the driver of a runaway tram which he can only steer from one narrow track on to another; five men are working on one track and one man on the other; anyone on the track he enters is bound to be killed". Philippa Foot, "The Problem of Abortion and the Doctrine of Double Effect," in Philippa Foot (ed.) *Virtues and Vices and Other Essays in Moral Philosophy* (Oxford: Basil Blackwell, 1978), p. 23 [19-32]. Originally published as Philippa Foot, "The

Whatever else we want to say about the success of these thought experiments, they all countenance scenarios far different from our experience. Of them, Singer's is really the only one that is even empirically *plausible*; the other two are almost farcical. Again, my contention is that these cases instantiate moral features with which we do have familiarity, but let us set that aside for now. If the problem with ticking time-bomb cases is that they outstrip our experience, then all of these thought experiments will surely have to be jettisoned as well, along with countless others. And, depending on what we mean by 'experience', so will hypotheticals such as "imagine you come home next Tuesday, and [whatever]"; even such a simple hypothetical outstrips our experience, at least in a trivial way. Above I proposed that the difference might have to do with morally relevant and morally irrelevant features, but I do not see what the argument would be to include these hypotheticals and yet exclude the aforementioned thought experiments.

Regardless, it seems that there are various possibilities. The first is that ticking time-bomb cases are excluded from our moral thinking, but violinists, shallow ponds, and trolleys get to stay. This seems implausible to me, but more will be said about it below. The second is that *everything* that even minimally outstrips our experience gets jettisoned, including simple hypotheticals. This cannot work as it would completely impoverish our moral thinking. The third proposal would be some sort of compromise wherein simple hypotheticals get to stay, but ticking time-bombs, violinists, shallow ponds, and trolleys are all out. I suspect that this is the sort of idea that Davis has in mind, but I want to saddle his critique of ticking time-bomb cases with an indictment of these other cases as well. And, while not as impoverishing as the second proposal, this third one still substantially impoverishes moral philosophy. Leaving aside ticking time-bombs for now, violinists, shallow ponds, and trolleys have played critical

---

Problem of Abortion and the Doctrine of Double Effect," *Oxford Review* 5 (1967), p. 5-15. Note that, in Foot's case, the driver has to choose which track to steer the trolley onto whereas, in Thomson's retelling, the trolley is already going toward the five and the question is whether to redirect toward the one. None of these details matters for present purposes.

roles in helping us to think about abortion, our duties to the poor, and the doctrine of double effect; without these constructs, those inquiries would be much worse off. And this is true even if the thought experiments fail as those failures are important for catalyzing discussion and developing alternative ideas. The same holds true for ticking time-bomb cases.

My reply to Davis—or anyone else of similar spirit—therefore has various prongs. First, I deny that ticking time-bomb cases outstrip our experience in any morally relevant way. Second, even if they did, this criterion threatens to undermine a broad and important swath of moral philosophy, perhaps even crippling the discipline. Third, even if that challenge can somehow be mitigated, various other useful thought experiments would be in at least as much trouble as ticking time-bomb cases; moral philosophy would be worse off without these. The costs of jettisoning these cases is too high, especially since nothing even stands to be gained. Therefore, we should keep these thought experiments as part of our moral methodology. Of course none of this is to say anything about particular normative conclusions, rather it is just to allow esoteric thought experiments as part of our philosophical toolkits.[36]

Let us now turn to Davis's third objection. It will be useful to say more than is minimally required in response here given that this objection raises issues to which we will return in §§8.3-8.4. Davis thinks that we are intuiting "an excuse rather than a justification"[37] when we reflect upon torture in ticking time-bomb cases. This thought is motivated by something that Shue wrote: "[if] the situation approximates those in the imaginary examples in which torture seems possible to justify, a judge can surely be expected to suspend the sentence."[38] Davis thinks that this indicates that we find the torturer

---

[36] Of course, I acknowledge that more could be said in about the role of thought experiments in moral methodology, though that takes us too far afield of the current project. This is an area that has historically received less discussion than would be expected, though I suspect that the newfound interest in experimental philosophy—with its strong reliance on thought experiments—portends a corrective. For a more general discussion of thought experiments in moral philosophy, see Jonathan Dancy, "The Role of Imaginary Cases in Ethics," *Pacific Philosophical Quarterly* 66 (1985): 141-153. For a discussion of the role of thought experiments in philosophy more generally, see, Roy Sorenson, *Thought Experiments* (New York: Oxford University Press, 1998).
[37] Davis (2005), p. 161.
[38] Shue (1978), p. 143. Cited in Davis (2005), p. 172.

"[g]uilty but excused", which "suggests doubts about moral justification".[39] These various comments invite us to think about an important point: what is it that we are meant to intuit in ticking time-bomb cases? As I said above, the answer needs to be moral permissibility—or justification, this difference is not important here—for ticking time-bomb methodology to work. If we are intuiting excuse rather than justification, that would indeed be a problem for this methodology.

So, first, what is the difference between an excuse and a justification? Or even before that, let us start with a suspended sentence. Suspended sentences have nothing to do with either excuses or justification: a suspended sentence is a form of judicial discretion that can be issued to show leniency, often to first-time offenders or as part of a plea bargain. Shue suggests that a suspended sentence might be appropriate for someone who tortured under ticking time-bomb-like conditions. But a suspended sentence means that such a person would have to be *convicted*, and I do not think that the ticking time-bomb torturer should be convicted. Again, we will take up that issue Chapter 8; for now, back to Davis. Justification and excuse are means to *acquittal*: a suspended sentence comes after an ascription of guilt, which is exactly what justification and excuse are meant to forestall. In other words, Davis's "guilty but excused" is not a proper category since excuse precludes guilt (cf., *not* guilty by reason of insanity).

To see this, let us now return to the difference between justification and excuse. While these concepts are important hallmarks of our criminal law—and moral thinking—they are somewhat more complicated than might be expected.[40] Consider, for example, Joshua Dressler: "a justification claim…seeks to show that the act was not wrongful, an excuse…tries to show that the actor is not

---

[39] Davis (1975), p. 172.
[40] For more discussion, see Marcia Baron, "Justifications and Excuses," *Ohio State Journal of Criminal Law* 2 (2005): 387-500. See also Douglas Husak, "On the Supposed Priority of Justification to Excuse," *Law and Philosophy* 24 (2005): 557-594. I thank Hans Allhoff and Matt Hanser for helpful discussion on differences between justification and excuse.

morally culpable for his wrongful conduct."[41] Or Sharon Byrd: "an excuse…not only presupposes the violation of a legal or moral norm, but also the wrongful or unjustified nature of this violation."[42] Or, finally, Michael Moore: "a justification shows that *prima facie* wrongful and unlawful conduct is not wrongful or unlawful at all…by contrast, an excuse does not take away our prima facie judgment that an act is wrongful and unlawful; rather, it shows that the actor was not culpable in his doing of an admittedly wrongful and unlawful act."[43]

One of the principal differences between justification and excuse is that, if someone is justified, then he did not do anything wrong; if he is (merely) excused, then he did something wrong, but it is not his fault. Self-defense and necessity are justifications: when these are adequately established, we acknowledge that the accused did not do anything wrong.[44] Excuse, though, goes to incapacity, such as would be manifest through duress or insanity. If the accused kills a family but can establish insanity, then we do not hold him (morally or criminally) liable since it was not his fault. But, in excusing him from legal punishment, we do not say that the killings were justified.

To further elucidate this distinction, consider that justifications attach to acts, whereas excuses attach to actors.[45] If, for example, someone were justified in killing one person to save five others (cf., the necessity defense; see §8.3), then some other person would have been also. In other words, what matters is the act of preventing the greater harm, not who perpetuates that act. Excuses, on the other hand, are different: the incapacities that they portend—like duress or insanity—are personal and attach to individuals. Second, claims of justification rest on a balancing of interests and considerations of the

---

[41] Joshua Dressler, *Understanding Criminal Law*, 3rd ed. (New York: Lexis Publishing, 2001), p. 558. Quoted in Husak (2005), p. 558.

[42] B. Sharon Byrd, "Wrongdoing and Attribution: Implications beyond the Justification-Excuse Distinction," *Wayne Law Review* 33.4 (1987), pp. 1289-1290 [1289-1342]. Quoted in Husak (2005), p. 558-559.

[43] Michael Moore, *Placing Blame* (Oxford: Clarendon Press, 1997), p. 674. Quoted in Husak (2005), p. 559.

[44] While I take this to be uncontroversial, see David Cohen, "The Development of the Modern Law of Necessity: A Comparative Critique", *Journal for History of Law* 4 (1985): 215-234. (Cohen argues that the German conception of necessity is importantly different from that indoctrinated in Ango-American law.)

[45] Eugene R. Milhizer, "Justification and Excuse: What They Were, What They Are, and What They Ought to Be," *St. John's Law Review* 78.4 (2004), p. 726 [725-895].

greater good, while excuses do not. Killing one to save five could be justified, but killing however many could still be excused if the killer lacked capacity (i.e., the numbers do not matter). Third, legally speaking, justifications acknowledge exceptions to prohibitions on criminal law, whereas excuses do not.[46] While the distinction between justification and excuse can be complicated, blurred, or attenuated in practice,[47] this basic presentation is sufficient for our purposes.

Let us now return to ticking time-bomb cases: which is it that we are meant to intuit, justification or excuse? Contra Davis, the answer is unequivocally justification. Remember that excuse has to do with incapacity, but that is not at stake with the ticking time-bomb torturer. Return to our above three criteria. First, our moral evaluation of the torture would not change whether the torturer was one person or another; the assessment is on the act of averting the egregious harm. Second, balancing of interests is precisely what is it stake and, should that balance not come out properly, then the torture was inappropriate. And, third, to endorse torture in those cases is to acknowledge that a universal prohibition on torture is misguided (if expedient). It bears emphasis that the most plausible legal justification for torture is necessity; we will return to this discussion §8.3.

The goals of this section have been to articulate and defend some of the central features of ticking time-bomb methodology, as well as to consider some objections to that methodology. As I proposed earlier, we should separate these methodological issues from the empirical ones: ticking time-bomb cases do not make empirical claims, so empirical objections miss the point. Those objections loom when we try to figure out how to apply ticking time-bomb cases to the real world, but we have various stops before getting to that project. For the rest of this chapter, I propose to accept ticking time-bomb methodology as philosophically legitimate and to further explore some of the central features of these cases.

---

[46] American Law and Legal Information, "Justification and Excuse: Similarities and Differences," *Crime and Justice*, Vol. 2. Available at http://law.jrank.org/pages/1114/Excuse-Theory-Justification-excuse-similarities-differences.html (accessed June 4, 2009).

[47] See, for example, Milhizer (2004). See also Baron (2005).

5.3   Ticking Time-Bomb Case Variants

If ticking time-bomb methodology is defensible, then we can look more closely at the details and logic of ticking time-bomb cases. As discussed in §5.1, at least two features of these cases are (allegedly) conspicuous: their dependence on utilitarian thinking and their high degree of idealization and abstraction. In that discussion, I did not say much about what idealization and abstraction were, so let us briefly consider those here. I take these terms from a recent essay by Shue; in this essay, Shue complains about the ticking time-bomb methodology that I defended in §5.2. In doing so, he invokes a distinction between idealization and abstraction which owes to Onora O'Neill:[48]

> Why are imaginary examples like ticking-bomb hypotheticals so badly misleading about how to plan for real cases? They mislead in two different ways that compound the error: idealization and abstraction. Idealization is the addition of positive features to an example in order to make the example better than reality, which lacks those features. Abstraction is the deletion of negative features of reality from an example in order to make the example still better than reality. Idealization adds sparkle, abstraction removes dirt.[49]

As examples of idealization, Shue points out various assumptions built into ticking time-bomb cases: that we have apprehended the right person, that torture will lead to prompt and accurate disclosure, and that torture will be practiced in rare and isolated cases.[50] There are actually even more idealizations than these, some of which I discussed in §5.1. For example, it is also implicit that, not only will we get the right information from the right person, but that the information will actually be used to disarm the threat. The bomb squad is infallible, time constraints are superable, the bomb is readily accessible, and so on. Shue's principal abstractive feature has to do with the institutional features that torture requires: he thinks that we simply could not have (effective) torture—even in one-off cases—without various institutions supporting it. And, following from this, Shue is skeptical about whether we could have one-

---

[48] Onora O'Neill, "Ethical Reasoning and Ideological Pluralism," *Ethics* 98 (1988), pp. 705, 711-712 [705-722]. Referenced in Shue (2006), p. 231.
[49] Shue (2006), p. 231.
[50] *Ibid*., p. 233.

off torture at all, that torture could be "conducted by wise, self-restrained angels."[51] Rather, he thinks that torture is inherently subject to abuse and nefarious spread.

Unfortunately for me, this all sounds quite compelling. Nevertheless, it runs together various issues that can be usefully separated. First, Shue thinks that ticking time-bomb cases need to be reasonably similar to real-world cases in order for the former to inform the latter; I am dubious about this. Regardless, we will talk more about torture and the real world in the third part of the book. (I should also say that I am more sanguine than Shue about how empirically plausible ticking time-bomb circumstances are.) Second, Shue's claims about abstraction make substantive claims about torture that I simply think are false: I think that torture can be (successfully) practiced without extensive institutional structures and, furthermore, that safeguards could be put in place to prevent the abuses that he worries about. For present purposes, we just need to agree that these are open questions; they will receive more discussion in subsequent chapters.

This then leaves idealizations, and I actually am sympathetic to the idea that ticking time-bomb cases are highly idealized, both in the ways that Shue suggests and in the others that I mentioned. But there are two things to say here. First, whether they are idealized or not has to do with empirical circumstances that we have yet to consider. One cannot simply label the cases as idealized without looking at actual cases. Presumably the critic agrees which gives rise to his contention that, by looking at actual cases, we will realize how idealized ticking time-bomb cases are. But I worry that this gets the quantifiers wrong: surely not *all* actual cases are like ticking time-bomb cases, but nobody ever said they were. If the ticking time-bomb proponent could even come up with a *single* real-world case that bears the alleged idealizations of ticking time-bomb cases, then ticking time-bomb methodology is at least vindicated in that case.

---

[51] *Ibid*., p. 235.

But, second, let us grant that ticking time-bomb cases are idealized. Why would idealization make those cases less philosophically useful? The concern is that ticking time-bomb methodology elicits intuitions about some cases (viz., the idealized ones) and then presumes to transfer those intuitions uncritically to different cases (viz., non-idealized cases). Or, in other words, ticking time-bomb cases elicit different intuitions than less-idealized variants and, therefore, these former intuitions are not useful in thinking about the latter cases. I actually suspect that this is false in both regards, but the former is a claim that we can test experimentally. If it can be shown to be false, then the worry about idealizations can be forestalled.

We can further use the experimental approach to gather data about the relationship between intuitions in ticking time-bomb cases and simple utilitarian thinking. As we will see, some of the critics of ticking time-bomb cases have alleged that such cases codify some sort of naïve utilitarianism and, as they quickly reject utilitarianism, they can similarly reject ticking time-bomb methodology. I think that the intuitions ticking time-bomb cases elicit are actually more subtle than this and include important non-utilitarian considerations; if this is correct, then the anti-utilitarian critique of ticking time-bomb methodology fails. We can gather empirical data about intuitions regarding ticking time-bomb cases and their putative utilitarian bases that will help us to adjudicate this issue.

To test these ideas, I designed four thought experiments, which varied along two dimensions; these are simple variants of the standard ticking time-bomb cases, such that two variables are exploited.[52] In standard cases, the torture is effected against a *terrorist* who is *guilty* of perpetuating some terrorist plot. The first variable, *guilt/innocence*, trades on this feature by proposing torture on an *innocent* (*viz.*, the terrorist's daughter[53]) with the effects otherwise being the same (*viz.*, the terrorist

---

[52] I thank Joshua Knobe for helpful discussions about the design of the cases.

[53] When this research was presented, some objected to the use of 'daughter' rather than the more gender-neutral 'child'. I am not sure why this matters. If the sense is that torturing daughters is worse than torturing children (who may be sons), then presumably there is a reason why. For example, if daughters are thought to be especially innocent (and therefore especially undeserving of torture), then this is precisely the most appropriate wording

giving up the location of the bomb). Also, in standard cases, the torture is *certain* to lead to the saving

of many people: as mentioned in §5.1, we are certain that the terrorist possesses material knowledge

to the location of the bomb; that the torture will generate the release of that information; and that the

information will lead to the timely deactivation of the bomb. As against this certainty, we might

postulate *uncertainty*: these cases offer a 1% chance of saving the lives rather than certain saving in

their converses, giving us the second variable, *(un)certainty*.[54]

It is worth offering two notes on the cases with uncertainty. First, these cases are silent as to

why there is uncertainty: maybe the terrorist does not know the location of the bomb; maybe the

torture will not work; maybe the information will not lead to the deactivation of the bomb. The

uncertainty cases can accommodate any of these sources of the uncertainty since any of them could

lead to a reduced likelihood of saving lives. Second, in the uncertainty cases, the number of lives saved

(10,000) is higher than the number of lives saved in the certainty cases (100). This is to keep the

expected numbers of lived saved through torture to be the same since the lower likelihood of success in

the uncertainty cases (1%) as against the certainty cases (100%) renders the expected outcomes

identical. Were the outcomes not to be identical, it would not be clear whether the (un)certainty or the

---

given the features that I want to test. However, maybe the objection is supposed to be that there is some irrelevant emotional feature that attaches to daughters, thus distorting the results.

Regardless, given that the objection came up multiple times, I re-ran the experiments with a change in wording: the responses with 'daughter' were compared to responses with 'child'. The analysis was done using a post-hoc linear contrast: $F_{(2, 445)} = 0.40$, $p = 0.52$. Qualitatively, this means that there was no statistical difference between moral judgments in cases using 'daughter' and those using 'child'.

[54] The one sort of uncertainty that this methodology cannot accommodate is uncertainty about whether we have actually apprehended a terrorist. As will become clearer shortly, I need four different cases for my aims: guilt/certainty, guilty/uncertainty, innocence/certainty, and innocence/uncertainty. Using 'suspected terrorist' rather than 'terrorist' accommodates further uncertainty, but it also rules out this second case since the suspected terrorist is not necessarily guilty.

The point of these thought experiments is therefore not to remove *all* of the idealizations from ticking time-bomb cases. For example, they also do not test the rare, isolated case idealization that Shue postulates (though which I think is less empirically suspect than some of the other alleged idealizations). Rather, the goal is to remove as much idealization as is straightforwardly possible and then to make some observations about the results, including as to how they might extend.

number of lives alone were driving the moral judgments, which would then preclude the proper analysis.[55]

As already indicated, we therefore have two variables—guilt/innocence and (un)certainty—which can be combined in the following ways:

	Guilt	Innocence
Certainty	Case 1	Case 3
Uncertainty	Case 2	Case 4

Here are the cases, along with the prompts that were included on the surveys:

Case 1: Guilt/Certainty. Imagine that you have just apprehended a terrorist who is responsible for planting a bomb in a crowded metropolitan center. The bomb squad has been unable to defuse the bomb and, unless the terrorist provides the deactivation code, it will detonate and kill 100 people.

You have exhausted all other possibilities and must now contemplate more extreme measures. If the terrorist is subjected to moderate torture, then he will surely provide the deactivation code for the bomb in time for its safe deactivation.

It is morally permissible to torture the terrorist.

Case 2: Guilt/Uncertainty. Imagine that you have just apprehended a terrorist who is responsible for planting a bomb in a crowded metropolitan center. The bomb squad has been unable to defuse the bomb and, unless the terrorist provides the deactivation code, it will detonate and kill 10,000 people.

You have exhausted all other possibilities and must now contemplate more extreme measures. If the terrorist is subjected to moderate torture, there is a 1% chance that he will provide the deactivation code for the bomb in time for its safe deactivation. However, there is a 99% chance that the torture will accomplish nothing and that all the lives will be lost.

It is morally permissible to torture the terrorist.

Case 3: Innocence/Certainty. Imagine that you have just apprehended a terrorist who is responsible for planting a bomb in a crowded metropolitan center. The bomb squad has been unable to defuse the bomb and, unless the terrorist provides the deactivation code, it will detonate and kill 100 people.

---

[55] Some critics have objected to varying the number of lives on the grounds that doing so would affect the results despite the associated percentages. Having already given the reasons for the adopted strategy, I allow that this is possible, even if there is no reason to think that it is true; more data in this regard is certainly invited. (There are various ways in which such a result would make no difference to my ultimate conclusions, though I shall not pursue that discussion here.)

You have exhausted all other possibilities and must now contemplate more extreme measures. The terrorist has been trained to resist torture, so torturing him to get the code will not be effective. You have learned the location of the terrorist's young daughter, who is completely innocent and knows nothing of her father's terrorist activities. Psychological profiling of the terrorist reveals that, if his daughter is subject to moderate torture, then he will surely provide the deactivation code for the bomb in time for its safe deactivation.

It is morally permissible to torture the terrorist's daughter.

Case 4: Innocence/Uncertainty. Imagine that you have just apprehended a terrorist who is responsible for planting a bomb in a crowded metropolitan center. The bomb squad has been unable to defuse the bomb and, unless the terrorist provides the deactivation code, it will detonate and kill 10,000 people.

You have exhausted all other possibilities and must now contemplate more extreme measures. The terrorist has been trained to resist torture, so torturing him to get the code will not be effective. You have learned the location of the terrorist's young daughter, who is completely innocent and knows nothing of her father's terrorist activities. Psychological profiling of the terrorist reveals that, if his daughter is subject to moderate torture, there is a 1% chance that he will provide the deactivation code for the bomb in time for its safe deactivation. However, there is a 99% chance that the torture will accomplish nothing and that all the lives will be lost.

It is morally permissible to torture the terrorist's daughter.

These cases were administered to 833 students, all of whom consented to participate in the research project. In order to preclude order effects,[56] each subject randomly received one of the cases; therefore, there were just over 200 responses to each case. The response was on a seven-point Likert scale, the standard sort of scale for such questionnaires.[57] On this scale, they were asked to indicate the degree to which they disagreed or agreed with the sentence that concludes the case; "strong disagreement" was indicated by '1' and "strong agreement" by '7'. Subjects were further asked to report their gender.[58]

A 2 x 2 ANOVA, with guilt/innocence and (un)certainty as the between-participant variables, revealed a significant independent effect for guilt/innocence, $F(1,829) = 90.9$, $p < .001$, $\eta^2 = .10$, but not

---

[56] In other words, the idea was to not have reactions to one case influence reactions to another case; had the research subjects seen multiple cases, it is possible/probable that their responses would have been biased as against subjects who only received a single case.

[57] For more discussion of Likert scales, see Rensis Likert, "A Technique for the Measurement of Attitudes," *Archives of Psychology* 140 (1932): 1-55

[58] They were also asked to report the number of philosophy classes that they had taken, including ones in which they were currently enrolled. Statistical analysis showed that this made no difference to the results, so this feature will be dropped from forthcoming discussion.

for (un)certainty, $F(1, 829) = 2.1$, p = .14. In addition, there was no interaction between the two variables, $F(1, 829) = .90$, p = .34.[59] Here are the mean results for each case:
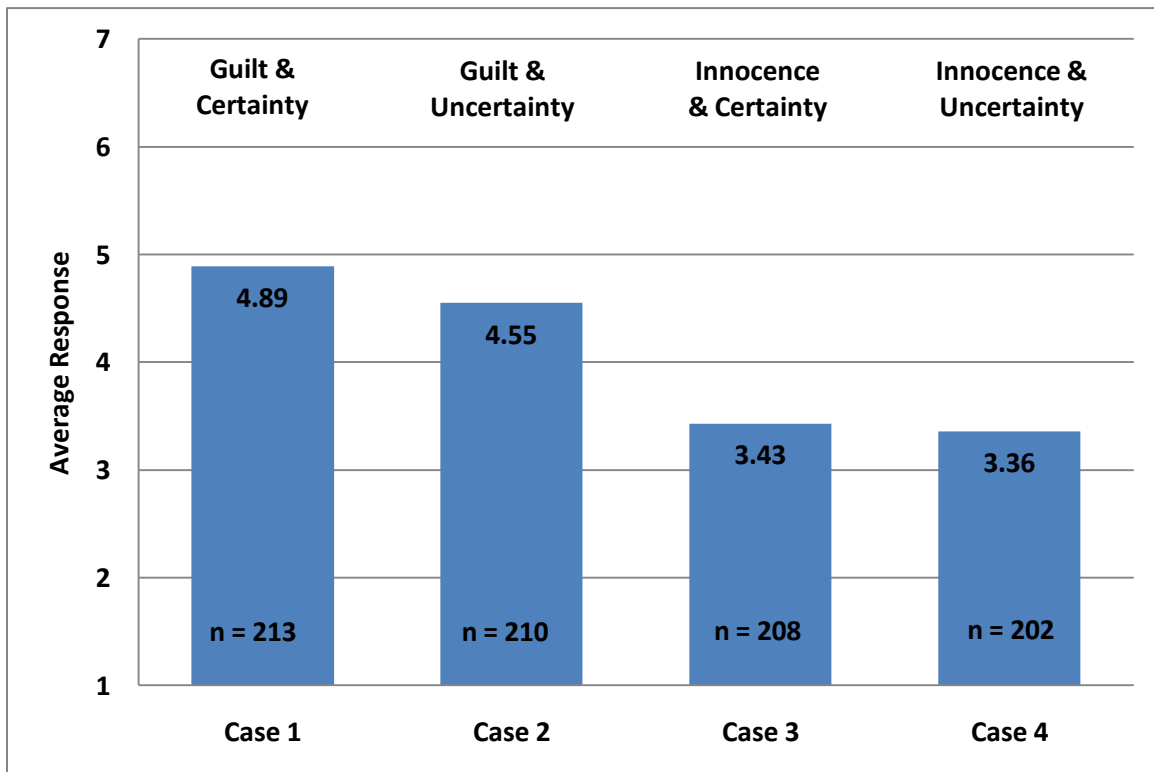

Figure 1: Mean Response by Case

Recall that, in Case 1 and Case 2, the torture would be effected against a *terrorist* and, in Case 3 and Case 4, it would be effected against the terrorist's *innocent daughter*. Here are the mean results for the guilt/innocence variable:
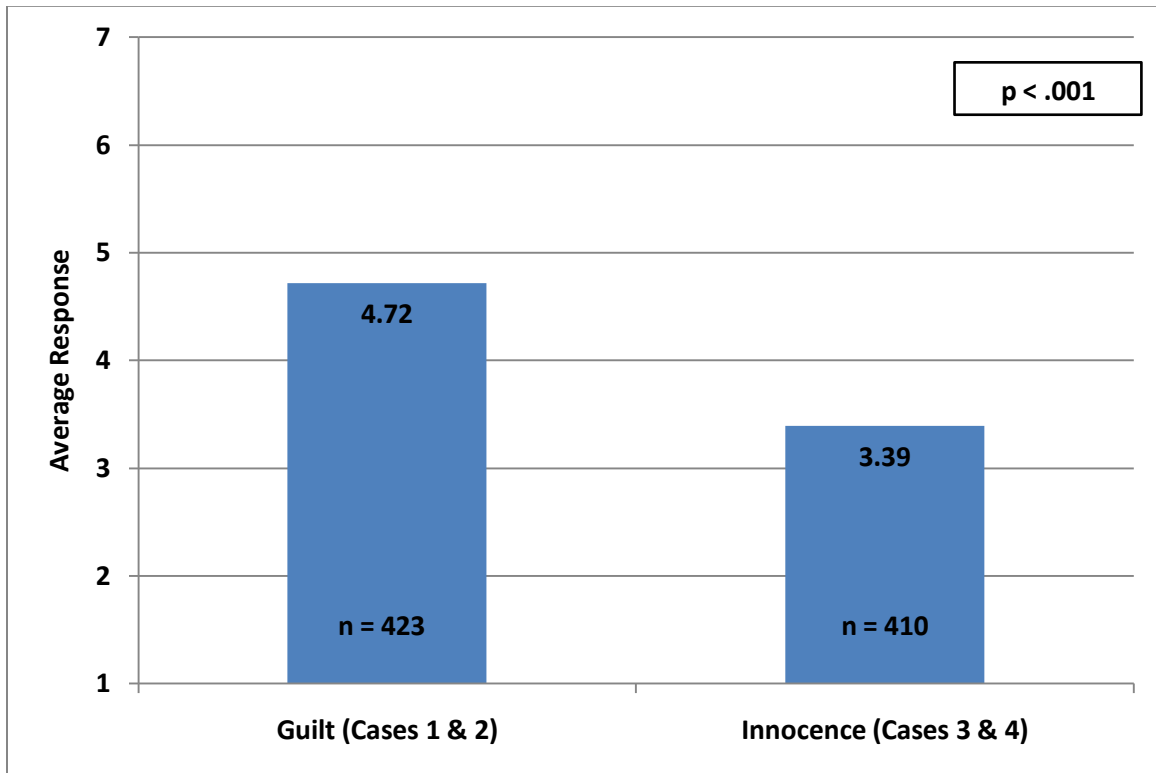

Figure 2: Guilt and Innocence Mean Responses

There is a substantial difference between the moral judgments in these categories and, indeed, one that is statistically significant.

The second variable, (un)certainty, can also be studied independently of the overall case averages. In Case 1 and Case 3, the torture was *certain* to save lives and, in Case 2 and Case 4, saving the lives was far less certain (only 1% likely). Here are the mean results for the (un)certainty variable:
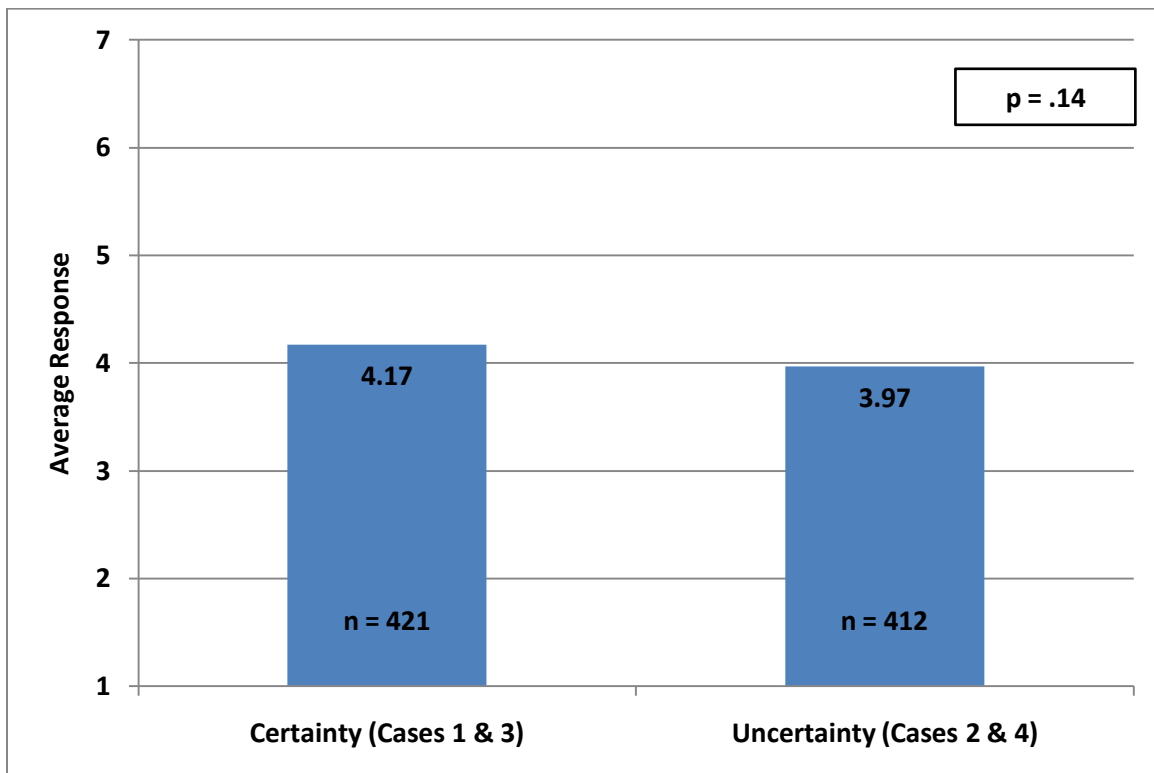
Figure 3: Certainty and Uncertainty Mean Responses

In this instance, there is also a difference in responses depending on whether the outcomes are (un)certain, but the response is *not* statistically significant.

Both of these results—i.e., guilt/innocence and (un)certainty—have direct implications for the use of ticking time-bomb cases. First, ticking time-bomb cases do not trade wholly on utilitarian considerations as is made clear by the guilt/innocence axis: it *matters* whether the subject of torture is a guilty terrorist or his innocent daughter. In each of these cases, the hedonic calculus comes out the same, yet the moral judgments are relevantly different.[60] This large literature, then, which focuses on

---

[60] Someone might argue that the consequences *could* be different as if, for example, word got out that we were torturing innocents; this might have negative societal effects. As the cases are constructed, however, such negative effects are not in play, and the wording of these cases was otherwise kept simple for obvious reasons.

the negative consequences of torture, misunderstands at least one central feature of the cases: the fact that we would torture a guilty terrorist is of primary moral importance.[61] Since guilt is a deontic notion, the cases invoke non-utilitarian features and criticisms that direct their animus toward utilitarianism are missing at least part of the story. Furthermore, notice that the critics of these cases have either criticized the cases' alleged utilitarian underpinnings as morally myopic (cf., invocations of dignity, inhumanity, rights, etc.[62]) or else have tried to show that the conclusions do not follow on utilitarian grounds alone; this was mentioned above and will bear more extensive treatment in the next chapter. Either of these criticisms, though, is compromised if non-utilitarian features help to drive our intuitions, which they do.

The implications of the (un)certainty results are less-straightforward, but my contention is that they mitigate worries about the idealizations. As already discussed, traditional ticking time-bomb cases suppose us to have various epistemic certainties, many of which we might (usually) expect to lack in the real world. Rather, the situation is far more likely to be that we have apprehended someone who: may or may not be a terrorist; may or may not know the location of the bomb; may or may not give up (correct) information during torture, and so on. Furthermore, such information may or may not lead to the saving of lives. In the cases designed to reflect uncertainty, there was a 1% chance that the torture would ultimately lead to saving lives. But, and this is critically important, there was no statistically significant difference in responses *regardless* of whether the outcomes were certain or uncertain. The certainty in standard ticking time-bomb cases, then, has not been shown to be psychologically efficacious regarding judgments as to the permissibility of torture. Or, to put it another way, the idealizations that these cases typically were not shown to affect the responses that we have to them.

---

[61] See, for example, Arrigo (2004), Bufacchi and Arrigo (2006), Luban (2005), Shue (2006), Wolfendale (2006), and Wynia (2005). It should be noted that not all of these authors consider *only* the consequences of torture and, furthermore, insofar as they do, this is surely a dialectic maneuver (i.e., to beat the proponents' arguments on their own terms), rather than an endorsement of consequentialism.

[62] Again, see Davis (2005). See also Michael Davis, "Torture and the Inhumane," *Criminal Justice Ethics* 26 (2007): 29-43.

There is no evidence that our intuitions are getting "messed up" by idealizations since they turn out to, statistically speaking, be the same whether idealizations are present or not.

As a final point, consider that survey data also reflected the gender of the subjects, and here are the means analyzing gender as a covariate, $F(1, 828) = 10.6$, $p = .001$:
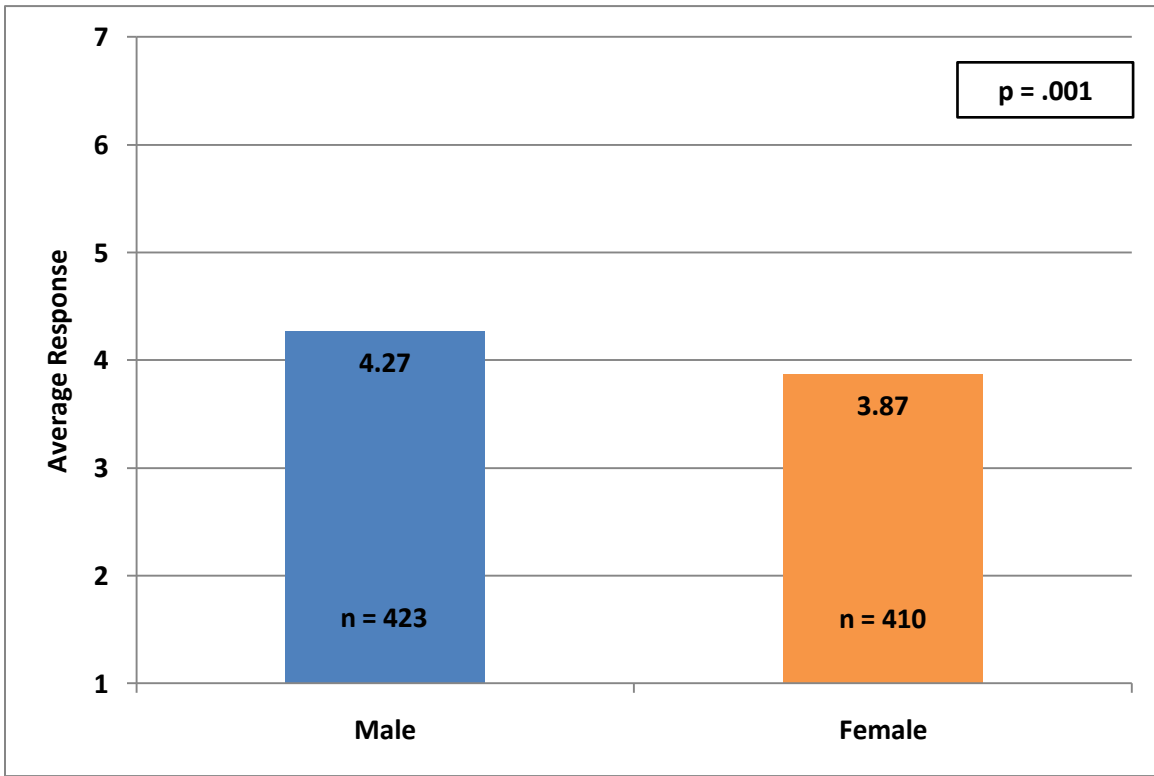

Figure 4: Mean Responses by Gender

Note, then that males were more likely to think that torture was permissible, and this difference was statistically significant (p<.001). Here are the mean responses for individual cases:
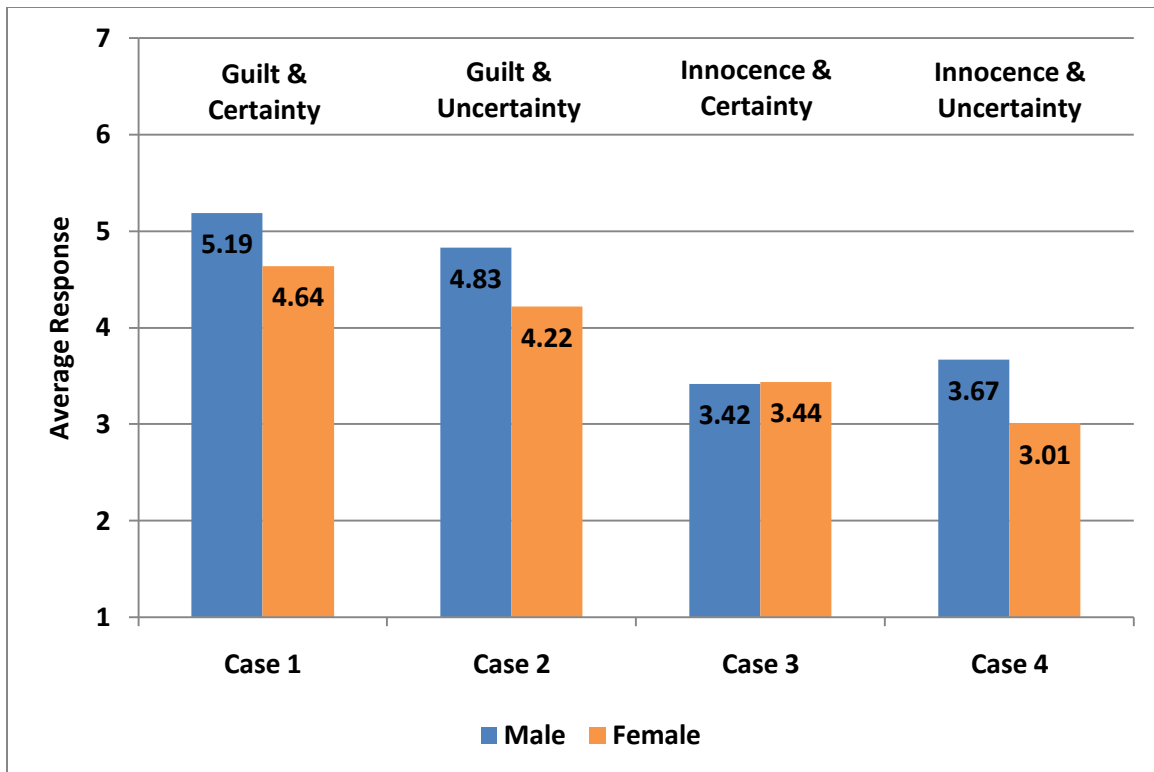
Figure 5: Mean Reponses by Case and Gender

In all but the third case, males, on average, were more likely than females to think that torture was morally permissible.[63] I do not want to lean on this result too heavily, but there have been studies that investigated differences in moral development and moral intuitions between men and women; part of this literature developed in response to Lawrence Kohlberg's work on stages of moral development.[64] Critics held that this account, which was undifferentiated across sexes, was incomplete since, even if there are insignificant sex differences in moral development, the sexes might differ in some aspects of

---

[63] Analyses were done using t-tests: Case 1: t(211) = 2.1, p = .039; Case 2: t(208) = 2.2, p = .03; Case 3: t(206)= .06, p = .95; and Case 4: t(200) = 2.3, p = .022.

[64] See, for example, Lawrence Kohlberg, "Moral Stages and Moralization: The Cognitive-Development Approach," in Thomas Lickona (ed.), *Moral Development and Behavior: Theory, Research, and Social Issues* (New York: Holt, Rinehart & Winston, 1976), pp. 31-53. See also Lawrence Kohlberg, *Essays on Moral Development: Vol. 2. The Psychology of Moral Development* (San Francisco: Harper & Row, 1984).

moral reasoning.[65] For example, Carol Gilligan has argued that males and females differ in their "orientations": males typically have a justice/rights orientation whereas females have a care/response orientation.[66] In subsequent work, Anne Kolby and Kohlberg suggested that this care/response orientation would manifest utilitarian thinking,[67] though the response has been critical.[68]

Of course, what is interesting about our results is that the males are more likely to find torture permissible than females, both overall and in three of four cases. To the extent that Kolby and Kohlberg are right—i.e., that females are more likely to be utilitarians—then it would follow that, in these ticking time-bomb cases, (wholly) utilitarian considerations are *not* the only ones present in the thought experiments since men were more likely to find torture morally permissible. Again, I do not want to place too much weight on this analysis, but the differences in gender response were striking: if the links between gender and moral orientation hold up, this is another reason to think that ticking time-bomb cases incorporate non-utilitarian considerations. It is not necessarily clear what those considerations would be, though. If the guilt of the terrorist was driving the reactions, then we would expect the men to have a higher average response in Guilt/Certainty and Guilt/Uncertainty, which they did. But then

---

[65] See Lawrence J. Walker, "Sex Differences in the Development of Moral Reasoning: A Critical Review," *Child Development* 55.3 (1984): 677-691. See also Lawrence J. Walker, "A Longitudinal Study of Moral Reasoning," *Child Development* 60 (1989): 157-166.

[66] Carol Gilligan, *In a Different Voice: Psychological Theory and Women's Development* (Cambridge, MA: Harvard University Press, 1982). For a philosophical analysis of Gilligan's work, see also Chesire Calhoun, "Justice, Care, Gender Bias," *Journal of Philosophy* 85.9 (1988): 451-463. For more on the differences between Kohlberg and Gilligan, see Owen Flanagan and Kathryn Jackson, "Justice, Care, and Gender: The Kohlberg-Gilligan Debate Revisited," *Ethics* 97.3 (1987): 622-637.

[67] Anne Kolby and Lawrence Kolhberg, *The Measurement of Moral Judgment*, vols. 1-2 (New York: Cambridge University Press, 1987). Referenced in Walker 1989, p. 158. Congruencies between consequentialism and care-based ethics have also been argued for in Julia Driver, "Consequentialism and Feminist Ethics," *Hypatia* 20.4 (2005): 183-199.

[68] See, for example, Michael W. Pratt, Gail Golding, and William J. Hunter, "Does Morality Have a Gender?: Sex, Sex Role, and Moral Judgment across the Adult Lifespan," *Merrill-Palmer Quarterly* 30 (1984): 321-340. See also Walker (1984) and Lawrence J. Walker, "Experimental and Cognitive Sources of Moral Development in Adulthood," *Human Development* 29 (1986): 113-24. Finally, see William J. Friedman *et al.*, "Sex Differences in Moral Judgments?: A Test of Gilligan's Theory," *Psychology of Women Quarterly,* 11 (1987): 37-46.

Results in experimental philosophy have been varied about the relationship between gender and utilitarian thinking; for example, Greene *et al*. reported gender effects in some trolley case variants and not in others. See Joshua D. Greene *et al*., "Pushing Moral Buttons: The Interaction between Personal Force and Intention in Moral Judgment," *Cognition* 111.3 (2009), esp. §2.2, §4.2. See also Jennifer Zamzow and Shaun Nichols, "Variations in Ethical Intuitions," *Philosophical Issues* (forthcoming), esp. §II.

men and women should have the same responses in Innocence/Certainty and Innocence/Uncertainty, yet men were still more likely to support torture in Innocence/Uncertainty, if not Innocence/Certainty. Regardless, the stronger support for the non-utilitarian considerations is still the (un-gendered) Guilt/Innocence results.

This chapter has covered much ground, so it might be useful to tie together various themes that have been herein developed. In §5.1, we considered the origins of ticking time-bomb cases, as well as looked at some of the features that various formulations of these cases share. In that discussion, we saw the work of Jeremy Bentham, as well as the consonance between ticking time-bomb cases and classical utilitarian thinking. §5.2 considered various methodological objections to the use of ticking time-bomb cases. I defended the use of the cases against those methodological objections, arguing that they play a legitimate role in our moral theorizing.

§5.3 considered various variants of ticking time-bomb cases in the aims of reaching two substantive conclusions. First, I argued that our moral judgments in these cases are responsive to non-utilitarian features; this is important given the affinity Bentham's hedonic calculus bears to the cases as well as allegations that have come from critics. Second, I presented evidence that moral judgments in idealized ticking time-bomb cases are statistically indistinguishable from less-idealized cases. This result goes against objections that traditional ticking time-bomb cases are ill-formed since the idealizations were not shown to be cognitively efficacious. Therefore, it is simply not a strike against ticking time-bomb cases that they are idealized—if, in fact, they are.

As the title of this chapter indicates, the issues that we have been herein discussed were methodological, as opposed to normative. Ticking time-bomb cases play such a central role in the debate about interrogational torture that this methodology deserves extended consideration. And, since the use of the cases has come under fire, a defense was needed if we are to proceed with them in

our arsenal. It is my hope that I have delimited a legitimate purview for their incorporation, as well as articulated some of the bases of the cases. Having now offered this methodological foundation, let us turn to its normative upshot.