

This article was downloaded by: [University of California Santa Cruz]
On: 12 November 2012, At: 12:31
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH,
UK



Australasian Journal of Philosophy

Publication details, including instructions for authors
and subscription information:

<http://www.tandfonline.com/loi/rajp20>

The Role of Intuition in Some Ethically Hard Cases

Daniel Guevara ^a

^a University of California, Santa Cruz

Version of record first published: 25 Feb 2010.

To cite this article: Daniel Guevara (2011): The Role of Intuition in Some Ethically Hard Cases, *Australasian Journal of Philosophy*, 89:1, 149-167

To link to this article: <http://dx.doi.org/10.1080/00048401003610882>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

THE ROLE OF INTUITION IN SOME ETHICALLY HARD CASES

Daniel Guevara

Among the hardest cases in the ethics of killing are those in which one innocent person poses a lethal threat to another. I argue in favour of the intuition that lethal self-defence is permissible in these cases, despite the difficulties that some philosophers (e.g., Otsuka and McMahan) have raised about it. Philosophers writing in this area—including those sympathetic to the intuition (e.g. Thomson and Kamm)—have downplayed or ignored an essential and authoritative role for intuition *per se* (as against discursive general principles): one based in moral sensibility and imagination rather than discursive argument or conceptual analysis. I am concerned to call attention to and rehabilitate this role for intuition.

1. Introduction

An interesting and well known hard case in the ethics of killing is the one where a bad man has thrown someone from a balcony in order to kill someone else—say *you*—sitting on the deck below.¹ Luckily (?) for you, you can swing into place a strong metal awning in order to shield yourself from the falling victim. You see him coming, but you can't spring from your recliner in time to get out of the way and, since he is large, he will kill you unless you shield yourself with the awning. But hitting the awning will kill *him*, since it will deflect him onto the street, many floors below. Then, a main ethical difficulty about shielding yourself in this case—call it the Balcony Case—is that the falling man is innocent, like you.

In spite of this difficulty, many people feel confident that it is morally permissible for you to shield yourself from the falling man in order to save your own life (even if, as we will assume, he will survive the fall if you do not shield yourself).² But, given the strictness of the moral injunction against killing innocents, the honest casuistry around this case and related cases—cases of so-called Innocent Threats and the like³—continues to resist a generally agreed-upon principle that justifies this intuitive sense of the

¹This is a variation on Judith Thomson's case [1991: 287], which itself is a variation on a case originally concocted by Robert Nozick [1974: 34–5], so far as I know.

²I use 'ethical' and 'moral' (and the relevant derivatives) interchangeably.

³See Otsuka [1994] and McMahan [1994] for a fairly thorough consideration of relevant cases, including so-called Innocent Aggressor Cases, discussed below.

cases. In fact, some recent literature seems to show that it is morally *impermissible* on the most compelling principles [Otsuka 1994, McMahan 1994].

But the intuition is defensible independently of ethical principles, and even if contrary to the best of them. This is because its moral authority and justificatory force depend as much or more on certain features of our sensibility than on what discursive reasoning from the relevant moral concepts or principles might seem to imply. It will become clearer, eventually, what I mean by this contrast between sensibility and reason, although it might help to say right away that it has Kantian undertones and echoes some themes in the work of P. F. Strawson (see §7.2, below).

I should emphasize at the outset that although I go into some detail about the morality of self-defence against Innocent Threats and related cases, my main concern is with the primacy of intuition over general principles in these and, potentially, other cases. The intuition regarding the permissibility of killing the innocent in self-defence is the focus of investigation as a means of making the more general point that intuition can be grounded in something that can justify it independently of principle (and even when in conflict with the best principles).

Then, with this more general point about the primacy of intuition in mind, I will be concerned with the following claim: untutored and inarticulate as it may be, the intuition—that lethal self-defence against Innocent Threats, and related cases, is morally permissible (if also tragic and deeply regrettable)—has grounds peculiar to it that can justify it regardless of what principle may seem to say. I argue that this claim is plausible, and should occupy an interesting position in a dialectic that has ignored or overlooked it. A comprehensive presentation and defence must be left for another occasion; here I show merely that the particular grounds I propose for the relevant intuition provide a more reasonable basis for adhering to the intuition than is provided by some state-of-the-art results of the search for a principle.

2. Recent Work on Innocent Threats

I discuss, in particular, a principle advanced by Judith Jarvis Thomson [1991] and echoed by F. M. Kamm [1992: chapters 2–3], whose contributions are state-of-the-art results of the search for a deontologically grounded principle (or principles) to justify the intuition that lethal self-defence against Innocent Threats is permissible; I know of no better investigation into such a principle than these, and my interest here is not in consequentialist principles.

I assume throughout that the deontological principle—if there is one—that justifies the common intuition about Innocent Threats (and related cases) is a principle of the very general and systematically interconnected sort that philosophers typically maintain or develop in ethical theory (like, say, the principle of Doing and Allowing, or of Double Effect, both of which serve later to illustrate what I mean). This is the sort of principle I am

contrasting with intuition. It seems clear that Thomson and Kamm have not produced such a general and systematic principle, nor even a more modest one, because in fact what they propose does not much improve upon the relatively untutored and inarticulate intuition that it is permissible to kill Innocent Threats in self-defence. So, it is not that they have offered a principle which, though plausible and full of potential, experiences complications that one might expect of any principle, and that can perhaps be settled by reflective equilibrium or some such.⁴ Rather, they have made hardly any progress towards a principle of the sort we are typically after in ethical theory. Recent work by Jeff McMahan [1994] and especially Michael Otsuka [1994] helps show this. These authors make a forceful case for the *impermissibility* of lethal self-defence against Innocent Threats, and hence against the possibility of a principle showing it to be permissible. I try to distil and elaborate the essence of their arguments in my criticisms of Thomson and Kamm. But, at the same time, I develop and defend a reasonable basis for resisting the conclusion Otsuka draws, and McMahan is (reluctantly) drawn to, that the intuition in question is therefore unjustifiable and incorrect.

We can resist this conclusion thanks to an alternative that they, and almost everyone else in this controversy, have overlooked, ignored or downplayed: namely that our ethical sensibility can provide its own kind of support for our intuitive ethical judgments—a kind not necessarily dependent upon principles of the sort we often try to establish in philosophy, but one drawn from, for example, our emotional responses to imagining what it would be like to live contrary to those judgments. This alternative view of the grounds of ethical intuition makes an important point about the role of intuition in the ethics of hard cases (and probably beyond): namely that intuition does not necessarily derive its moral authority or justificatory force ultimately from principles arrived at through discursive reasoning, but rather from sources proper to intuition itself.

I cannot give here a full argument for this view of a distinctive ground and role for intuition in certain ethical cases. Instead, after presenting the arguments against the principle that Thomson and Kamm have suggested, I simply describe the view I am proposing, as plausibly as I can, with the idea that the relatively simple act of bringing the view out of its neglected quarters will win over or attract some to it immediately. I also block, if not rebut entirely, some main objections to my view, and show that philosophers like Thomson and Kamm especially ought to embrace it, given the role intuition plays in their intellectual lives already.

The harder thing is to develop the view into a satisfactory account of intuition in ethics generally. I make a start on this in my sketch of a connection to some Strawsonian themes.

⁴Although reflective equilibrium is central to many discussions of intuition and principle, it is not of direct concern to me here. For the intuition I am concerned with doesn't budge, as it were, in the light of what seems to be implied by compelling principles. I am proposing, therefore, another method and ground of defence for it, since I believe there is more to it than a certain kind of reflection on principles and cases can establish. (Norman Daniels' entry in the *Stanford Encyclopedia of Philosophy* [2008] provides excellent discussion of reflective equilibrium.)

3. Clarifying the Relevant Intuition

3.1 *The Casuistry*

Let me begin, now, to make my case by first clarifying the intuition that I mean to defend, which in turn will show what makes it hard to defend. I assume that those who permit the killing in self-defence in the Balcony Case also take seriously the strictness of the moral injunction against killing innocent people and accept a strong presumption against killing innocents even when this is the only way to save one's own innocent life. They accept standard cases reflecting this point: e.g. that we can't shoot to death our innocent neighbour—peacefully staring out of his high-rise apartment window across the street—just because we will be killed by someone else unless we do. The neighbour is an entirely innocent bystander, and many who feel confident about the permissibility of killing the falling man in the Balcony Case, or other Innocent Threats, feel just as sure of the *impermissibility*, or at least dubiousness, of killing the bystander. It is this intuitive sense of the cases that I wish to defend, even though it is very difficult and perhaps impossible to give a principled ethical reason for distinguishing the two killings.

It is difficult, for instance, to give a principle that allows us to construe the threat posed by the falling man as the sort of threat that justifies lethal self-defence against him, but that does not allow us to kill the neighbour who is peacefully minding his own business. After all, the falling man might have been just as peacefully minding his own business too, before being knocked unconscious and thrown off the balcony. There are very effective ways of sending these points home, as we will see later, after entertaining the relevant types of case in more detail.

As for details, my position will be compelling in the end only if it can be seen to emerge all right from the thicket of complications that the state-of-the-art casuistry of the cases has produced over the last few decades. So I must give explicit attention to the casuistry, even though this will include matters familiar to many readers.

We can begin to get a sense of the intricacies by elaborating on the Bystander Case. We can imagine that, e.g., a bad man has broken into our own high-rise apartment and will kill us unless we kill the innocent neighbour in the one across the street. Then, as in the Balcony Case, the whole terrible scenario originates with the bad actions of a bad man. This might seem to be a good way of keeping the cases as close as possible to one another. But it also introduces the following complication: the judgment that it is impermissible to kill the neighbour might be influenced by the thought that we ought not to give in to the evil threats of bad men. In the Balcony Case, as described, there is no issue of giving in or not to the bad man; so we've complicated the ethical question of interest by introducing issues around giving in or not to evil threats. So, instead of a bad man threatening us, we might imagine that we are threatened by someone blamelessly in a temporary fit of madness, or by an innocent but terribly confused child, or (if we may be science fictional) by a sophisticated security

robot experiencing an improbable glitch in its data or program, etc. There are other variations on the Bystander Case, which also avoid issues that can distract our attention from the intended point; but the variations just introduced are of a type that will interest us later when examining the concept of innocence.

In any event, the intended point is, of course, to question whether there could be a principled distinction that makes it permissible to kill the Innocent Threat, even though it is not permissible to kill the mere bystander (all other things more or less equal). I assume that such variations on Bystander Cases can get that point across; for all the complications that can and do arise in putting and discussing such cases, the similarities between Innocent Threats and mere bystanders are not too hard to see, and they are enough to raise questions about the intuition that, all things equal, it's permissible to kill the one but not the other.⁵

Let us return then to the similarities between mere bystanders and Innocent Threats. In the High-Rise Case, in order to eliminate the distracting issue of giving in to evil threats, we introduced another sort of innocent threat—a so-called Innocent Aggressor: e.g. a (blamelessly) temporarily insane lethal threat. We could just as well have introduced a temporarily drugged or sleepwalking lethal threat, or any other lethal agent who intends to kill you or is trying to kill you but who, because he is temporarily and blamelessly out of his right mind, cannot be held morally responsible for his lethal actions.⁶ Of course, this sort of innocence could be introduced into the Balcony Case too, by making the evil man there a temporarily insane man instead. Or, to take a slightly different tack, we could make him an inanimate object, like the sci-fi robot with a blameless glitch, or whatever. Or, as in some versions of the Balcony Case, we could eliminate all question of blame or innocence in the cause of the threat by imagining that the hapless man is blown off the balcony by a strong, tragic wind out of nowhere, or some such.

So, there are ways to keep the focus on the innocence of the Innocent Threat, in order to, as it were, test the purity of the enduring intuitive judgment that it is permissible to kill in self-defence some innocents (like

⁵So-called experimental philosophy raises a different sort of complication for anyone writing in my vein. Formal surveys seem to elicit intuitive judgments from the average person that are contrary to what philosophers often confidently assert to be intuitive or commonly held. Worse, these surveys seem to reveal an uncomfortable statistical significance between intuitive judgments and mere socio-economic status, ethnic or cultural background, and the like. There are many issues worthy of careful and critical treatment here (see Ludwig [2007]) and that experimental philosophers do not speak with one voice on (see Knobe and Nichols [2008]).

What's crucial to my thesis is that thoughtful people, who take seriously the presumption against killing innocents, have the relevant intuition, and that they continue to have it after initiation into the casuistry and relatively sophisticated philosophical reflection. It endures for them *in spite* of honest, critical reflection on principles that seem to challenge the intuition. So, of greatest importance to me are the enduring intuitions of people like Thomson and Kamm.

Will formal surveys show that the intuition I am interested in defending turns out to be (statistically) significantly related to mere cultural heritage, socio-economic status, gender, or training in a certain style of philosophy or the like? If so, what would this show about why people like Thomson and Kamm hold it, people who have made a career of thinking as carefully as possible about the issues, and who have subjected themselves to the greatest critical reflection, and who, for all we can tell, take very seriously the presumptions against killing innocents and so on? How would experimental techniques test the sort of limits in sensibility that I propose as the source of the moral authority of the enduring intuition in these hard cases?

These are among the many issues that deserve their own treatment in another place.

⁶See Otsuka [1994: 93] for why it may be important to stipulate that the madness is temporary.

Innocent Threats) and not others (like mere bystanders). So much, then, for our preliminary clarification of the intuition itself.

3.2 *The General Difficulty Presented by the Casuistry*

How clear is it that the intuition is especially hard to defend? It might appear to some that we are exaggerating the difficulties around being able to justify it with a principle. There are well-known and influential ethical principles that make it permissible in certain circumstances to kill, or at least do something that knowingly involves killing, an innocent. The Doctrine of Double Effect springs to mind. Roughly, it says that when killing an innocent is a foreseen consequence of what one does, it is sometimes permitted, so long as it is also an unintended side effect. This principle will serve later as a good example of the sort of general principle I have in mind when I speak of principles that philosophers typically seek in ethical theory, and in the struggle to resolve hard cases. But, whatever the virtues of a principle like Double Effect, it will not help with Innocent Threats: for the killing of an Innocent Threat (in the Balcony Case and related cases) cannot plausibly be conceived of in terms of the foreseen but *unintended side effect* of what one does. After all, you intentionally deflect the falling man in the Balcony Case, and his foreseen death is no more a side effect of your doing that than your neighbour's death would be if you intentionally pushed him out of his apartment window onto the street many floors below.

4. Criticism of a Principle Offered by Thomson and Kamm

4.1 *The Principle*

Let us turn, now, to the main literature. Judith Thomson [1991] has emphasized that an Innocent Threat, like the man about to smash you to death on your deck, is not exactly a mere bystander. He is, after all, about to kill you. So, as he falls towards you, he is a threat to your innocent life. As Thomson [1991: 288–9] says, he will kill you, unless you kill him.⁷ This is in contrast to the mere bystander—like your neighbour—who is not about to smash or shoot (or anything else) you to death, but is instead peacefully minding his own innocent business. Here, then, is a distinction that might support a principle to justify deadly force against Innocent Threats, while holding the line against bystanders. With this we might make a principled advance beyond relatively inarticulate appeals to intuition.

But if the measure of success in this controversy is the articulation of the sort of general principle that theorists typically seek, then I cannot in fact perceive the advantage over relatively inarticulate intuition. This is because the only principle that eventually emerges from the distinction depends on

⁷We are setting aside for the moment the important question of whether the *man* is the threat or just his body.

a fairly strong and questionable assumption: namely that a violation of your right not to be killed (or a violation of your right to the integrity of your body, or to your 'space', among other things) is threatened, or caused to be threatened, by the Innocent Threat [Thomson 1991: 301].⁸ This is an assumption made true, on Thomson's view of rights, by the fact that the Threat is about to kill you, whereas you are innocent in every way (and in particular, not at fault for the threat). This is a view of rights according to which a right can be violated even by someone who is not acting as an agent. The basic idea is put well by F. M. Kamm [1992: 47] when she says:

The position of [an Innocent Threat] is different from that of a natural object—for example, a stone that the wind hurls at a person—because she is not a stone, but, rather, a person who should not be in an inappropriate position relative to others. One person's inappropriate location *vis-à-vis* another raises moral questions no matter how it comes about, whereas the unfortunate location of an object does not One simply has a right not to have someone on the body or property to which one is entitled, even if the wind put them there.

4.2 The Criticism

That we have the right Kamm describes here, and that it extends to lethal defence against Innocent Threats, seems very plausible. But it is hard to see how this line of thought improves upon relatively inarticulate intuition, especially if the goal is the articulation of the sort of general principle that theorists typically seek in order to justify the relevant intuition. On the view proposed by Thomson and Kamm, if the Innocent Threat kills you, he *violates a serious right* (or causes the violation of a serious right)—or something relatively strong like that. This is what cancels out the moral protections such an innocent has in virtue of his many similarities to a mere bystander. But then the Threat violates such a right (or causes the violation of such a right) without doing or failing to do anything to violate it, and, therefore, without doing or failing to do anything to strip himself of the stringent moral protection against being killed that he has in virtue of being so innocent.⁹

We have been describing the falling man—rather than, say, his *body*—as posing a lethal threat. But, of course, given what I have just been saying, we must confront what's questionable and tendentious in this description, natural as it may be to say that the Threat will kill you unless you kill him. Thomson and Kamm are trying to apply the idea of a rights violation to someone who, *qua* agent, is no more involved in the violation than a falling stone, as Otsuka points out [1994: 80]. The application is evidently strained.

Now, if we believe that it *is* permissible to kill the innocent in the Balcony Case, then at least one philosophical incentive for applying the concept of a rights violation in this strained way is clear: we need something principled

⁸I add the parenthetical remark about the integrity of one's body etc., because Thomson's position is not necessarily limited to cases where the right to *life* is threatened, as Otsuka explains [1994: 82–3].

⁹[H]e is not doing anything at all—he is merely *falling* toward you', as Thomson says [1991: 287].

and strong, like a violation of a strict right to the threatened's life, to pit against the stringent protections that innocence presumably gives this particular threat.¹⁰ So one motivation for the view is understandable. Still, the view depends upon the idea of a rights violator (or causer of the violation of a right) who is not involved in a rights violation, except in the agent-less way that a stone might be (through gravity and the initial force of the throw). In fact, on Thomson's and Kamm's view, it is presumably irrelevant that there might have been agential involvement entirely in the opposite direction: e.g., that the Threat might have made every effort to *avoid* harming you in any way (by struggling against the evil villain, or the wind).

Otsuka is very effective in pressing to the hilt all the difficulties in this idea of an agent-less rights violation, an idea he thinks we must reject. But I want to make a different point. Whatever we might think about a rights violation that is credited to an agent even though it occurs without (and in spite of) her agency, we cannot simply take it to license a lethal response in self-defence against Innocent Threats, as though the violation were morally on a par with the uncontroversial sorts of violation: the sort that can be credited to an agent for, say, knowingly and intentionally doing (or failing to do) something to violate it.

All hands are agreed that the ethical difficulty from the start is due to the fact that innocence makes a difference to questions of self-defensive killing, as it does to killing or harming generally. The many similarities between Innocent Threats and mere bystanders bring this out in the cases at issue. There would hardly be any philosophical difficulty, initial or enduring, if we thought the bystander-like blamelessness of the killer didn't raise significant issues about moral parity and permissibility in defence against such a killer. These are issues everyone must face, even those of us who feel confident that it *is* permissible to kill Innocent Threats in self-defence. It is hard to say *why* it is permissible, given the strict presumption against killing innocents (even in the defence of one's own innocent life). But then these issues are hidden entirely from view when we reason simply and unqualifiedly as follows: He will kill me unless I kill him. I am innocent (and not threatening him, nor responsible for his threatening me, etc.) therefore, unless I kill him, he will violate my right not to be killed. Hence, he does not retain his right not to be killed. So it is morally permissible for me to kill him in self-defence.

This line of reasoning—essentially Thomson's [1991: 299–303]—would provide an elegant and principled solution to the case of Innocent Threats, including Innocent Aggressors (and much else perhaps), if it could sail smoothly through like that. It would constitute, in fact, a straightforward application of a pretty uncontroversial principle of self-defence as applied to lethal threats. But it conceals what is at issue, namely the fact that the threat he innocently poses is a threat fit for a falling stone. And innocence of this sort (blamelessness, faultlessness), or any other, always makes a difference

¹⁰In her more recent work, Kamm discusses cases that tend to show that the idea of a rights violation is less important than one might think [2007: 169]. But this adjustment does not help much with the reply to the line of objection one finds in Otsuka [1994] or McMahan [1994].

to the application of general principles in the ethics of harm, including, of course, the principle of self-defence.

The innocence of a threat makes a difference even when the agency of the threat *is* involved in the threat: e.g. in case of Innocent Aggressors, where the Innocent is trying to kill you, but temporarily insane. If the Aggressor succeeds, his particular sort of faultlessness or blamelessness is a mitigating factor in determining what we can do to him in response (e.g. in how to punish him, if at all). Likewise, if your awning fails to adjust, and I (the large man falling towards you) smash you and thus, cushioned by you, survive the fall, nobody thinks that I should be apprehended while it is decided what ought to be done to me for my having ‘violated’ your right not to be killed.

We may therefore recognize (as Otsuka does), with Thomson and Kamm, that the Threat will kill you unless you kill him. And, for the sake of argument, we may also allow (as Otsuka does not) that, since you are innocent (and non-threatening), the Threat will therefore *violate your right* not to be killed—even though this seems to strain the application of the terms just emphasized. We may likewise allow, for the sake of argument, that, as Kamm says, the inappropriateness of the Innocent Threat’s location and trajectory *vis-à-vis* the other innocent automatically raises moral issues. But none of this settles, on principle, that the Threat can be *killed* in self-defence. It would be like a bad play on words to say that it *is* settled on the principle that one has the right to kill in self-defence anyone who would violate one’s own right not to be killed. And, as I think Kamm would admit, it would not be plausible to say that one has the right to *kill* in self-defence *anyone* who has raised moral issues by being in a position or trajectory he should not be in with respect to oneself.

We haven’t advanced much, then, beyond the original, more or less inarticulate intuition that lethal self-defence in the Balcony Case is permissible. What’s lacking is an understanding of *how* the sort of threat involved overcomes, on principle, the strong presumptions set by innocence. We need this understanding if we are to make progress towards the desired principle. It seems safe to say that a plausible principle is not evident.

5. Double Effect and Doing and Allowing: General Principles of the Wished-For Type

Now there are widely accepted principles in this area of ethical concern (the ethics of killing, harming, and self-defence) that don’t, so far as I can tell, help us resolve the hard issues under discussion, but that at least approximate the sort of principle philosophers typically seek in order to justify, reject, explain, or organize our intuitions in a wide range of cases, including hard cases. Double Effect and Doing and Allowing are each good examples of this sort of principle. They are not universally accepted, nor thought to be entirely unproblematic even by those who do accept them, but they have enjoyed broad and fruitful engagement nevertheless, as even

convincingly critical accounts illustrate.¹¹ It is easy to see their potential significance for the ethics of killing (as for many other ethical issues): killing someone (innocent or not) seems obviously and importantly different from allowing her to be killed. Killing someone (innocent or not) as a foreseen, but unintended, side effect of some action seems to be too. Then, both principles employ from the start fairly uncontroversial ideas of agency and responsibility, with *prima facie* and natural application to a great variety of issues. These include intuitive, and intuitively related, ideas of responsibility, intention, knowledge, foresight, activity, and passivity, among others.¹²

Again, even able proponents of the principles find it difficult to give precise and entirely satisfactory formulations of them. But the most promising formulations command the philosophical respect and attention they do, *as principles*, or principled general distinctions, in large part because of how these just mentioned and other familiar ideas of agency and responsibility relate to each other and seem to guide often subtle and sophisticated application of the principles in a great variety of cases, including hard cases.¹³ Herein lies a broader account of why it is so difficult to give a deontologically grounded and principled defence of the intuition under discussion: the principles which already command general respect in the area seem to be informed by intuitions on the other side, namely, intuitions that tend to support strict moral protections for the innocent against intentional or foreseen harm, and harm that comes from active doing rather than passive allowing. These are protections from the harms attributable to responsible agency. And this is just one example of the centrality and weight of the notion of responsible agency in deontological morality, including the morality of self-defence. It is to Thomson's credit that she began long ago, with vivid and ingenious moral imagination, to question its centrality and weight,¹⁴ and to try to find principles that are more in tune with so much else of at least equal significance in deontological morality.

Still, there are also certain independent contributions of sensibility available to us. However, before I move on to those contributions, let me emphasize that I have focused only on Thomson and Kamm, and considerations consistent with deontological or non-consequentialist morality. There is therefore no question of my having shown that no other principle fares better than the one proposed by Thomson and echoed by Kamm. But I don't know of any principle that does; so I turn now to

¹¹See McIntyre [2001] for a demolition of, and Quinn [1989] and Kamm [2007] for perhaps the most subtle and sophisticated of sympathetic treatments of, Double Effect. Quinn [1989] includes discussion of its relationship to Doing and Allowing.

¹²The point is confirmed even by the best *unsympathetic* treatments. In McIntyre's [2001] critique of Double Effect, e.g., the familiar intuitions driving the principle are mostly not in question. What she argues is that it is very unlikely that we could formulate a principle, like the traditional principle of Double Effect (or any of its refinements), to cover the great variety of cases covered by the relevant intuitions and many refinements of the principle. More specifically, she argues that where Double Effect might seem to be doing the right work, it is superfluous, and (worse) can't do the right work in many cases where it is supposed to, thus supporting my thesis that intuitions have authority independent of an organizing, highly general principle.

¹³Kamm's ingenuity in illustrating this is renowned (e.g., Kamm [2007]). Cf. also Quinn [1989], especially for helpful discussion of the interconnectedness of the two principles.

¹⁴Beginning with her famous paper on abortion, and equally famous violinist case [Thomson 1971].

considerations which can support intuition directly and independently of the results of a search for a principle.

6. A Novel Defence of the Intuition in Question

What I would like to make plausible and interesting in the space that remains is this: given what lies behind it, the enduring, intuitive judgment of good and thoughtful people can itself serve as a justification of the permissibility of killing in self-defence in the Balcony Case, and other hard cases of Innocent Threats and the like. This is so even if it is impossible to produce a principle to decide the tough casuistical issues raised by and for this intuitive judgment.

What I have in mind can be brought out by certain questions or imaginative exercises. For example: could you imagine yourself, as a spectator in the actual circumstances, feeling the same way about the killing in Balcony Case as about the killing of a mere bystander, like the one in the high-rise? How would you feel about a society that treated the killing of Innocent Threats (or worse, Aggressors) on a legal par with the killing of mere innocent bystanders, e.g., came to treat them as murder?

I believe that these questions can evoke feelings that serve (perhaps together with other responses from sensibility) as the source for, and authority of, the intuitive judgment that it is permissible to defend oneself against Innocent Threats, all the similarities to bystanders notwithstanding. In this way, the questions evoke responses from moral sensibility that help make explicit the relevant intuition's significance and authority. There are other ways, more involved and developed, to evoke the relevant responses.¹⁵ But, however we do it, what's to be evoked, in the cases under consideration, is how we would feel about really having to live by the relevant judgment about Innocent Threats (or Aggressors): namely, the judgment that rules out lethal self-defence on principle because of the similarities between bystanders and Threats.

The point of calling attention to the fact that the judgment which allows lethal self-defence is the enduring intuitive judgment of good and thoughtful people is to indicate that the intuition has endured some thoughtful, critical reflection by people who take seriously the presumptions of innocence, and so forth, and generally have good judgment about the ethics of harm and killing (at least). The point of calling the judgment intuitive, and of urging that its origin and justification lie in sensibility rather than reason, is to call attention to the indispensable contribution of feeling to the validity of certain ethical judgments, a contribution in potential conflict with, and thus originating in something other than, the logical or analytical relationship of the relevant principles or concepts to each other (on impressive display in, for example, Otsuka's and McMahan's arguments).

So, as perhaps goes without saying by now, it would miss the point I wish to make if the evocative questions were taken to indicate the way to

¹⁵I discuss these in a longer version of this paper.

an underlying principle that favoured one response over another, and included therefore the response *I* favour: that one could *not* imagine feeling the same way about mere bystanders and Innocent Threats, nor less than repelled by a society that was committed to treating them as morally on a par. For one thing, the responses *I* favour might reflect consequentialist sensibilities as much as deontological ones. So they are exceedingly indeterminate in respect of what principles (if any) might account for them.

I am not concerned with consequentialist morality. So, the point is this: inasmuch as the feelings behind the responses *I* favour are those of reflective and conscientious people with deontological attitudes or sensibilities (like Thomson and Kamm, say), they can provide a perfectly reasonable backing for the judgment that killing in the Balcony Case is permissible. And they can do so without any further need for a deontologically-grounded principle that vouches for the feelings.¹⁶

It is important to bear in mind that this view is focused on hard cases, like Innocent Threats, even if it might have wider application. There is a special significance to disagreement in hard cases like those we are considering. For example, I think that (unlike moral disagreement in general) even if you do *not* have the intuition that killing in the Balcony Case is permissible, or even if you have the opposite intuition, you will likely find that you do not feel (and cannot easily imagine feeling) the same way about the two types of killing (Innocent Threat versus innocent bystander). And this is especially likely if you explicitly recognize something we all ought to recognize, namely that Innocent Threats make for hard cases in part *because* people like Thomson and Kamm—who have thought about all the issues as carefully as possible and who take very seriously the protections of innocence—have from the start the firm intuition that killing in the Balcony Case is permissible, and sincerely *continue* to have it in the face of the honest difficulties raised by the most sophisticated and analytically astute casuistry. Any fair attempt to make sense of the meaning of their intuition in these hard cases must keep all such things in view.

So, whatever our intuitions about the cases, we must recognize some significance in the fact that these philosophers, and others, persevere in the search for a principle to back the intuition, in spite of the casuistical difficulties. They persevere because they are confident they are showing good judgment about the cases. We have reason to think that their judgment is good in this general arena, and reason to try to conceive of their perseverance as something other than a sadly mistaken (or unconscious or self-deluded) obsessive search for a justification of murder or some such. Likewise, we must try to make sense of their persistent judgment in a way that steers between the horns of a certain kind of dilemma they seem to be stuck in, owing to their (questionable) assumption that their judgment in these cases requires a principle to validate it. As hopes dim for the discovery of the elusive principle, philosophers like Thomson or Kamm seem to be

¹⁶This is not to say the search for a principle is pointless or of little value. The opposite is true, for many reasons.

stuck with either (i) having to stubbornly and irrationally stand by consistently hard-to-justify mere gut and inarticulate intuition, or (ii) having to follow the discursive implications wherever they lead, without regard for the effects on sensibility. This is a false dilemma, because of the way it characterizes intuition. The characterization limits the possibilities for the ethical significance of an intuition to the eventual articulation of a principle (however elusive) of the wished-for type.

7. Objections and Replies

7.1 Objections Based on the Limited Authority of Intuition

Now good and thoughtful people (as good and thoughtful as one might expect anyone to be) have believed and done some indefensible (and worse) things, perhaps sincerely thinking they were showing good judgment.¹⁷ So the sort of authority I am relying on in certain intuitions is complicated by the fact that it isn't always easy to tell when the sincere intuitive judgments of even the best people are, actually, bad and thoughtless (or worse). But, then, the same can be true of their closely-reasoned arguments. And, in any case, there is a difference between (i) an ethical case, hard to judge correctly because so many good and thoughtful people are mistaken (self-deluded, etc.) about it, and (ii) a genuinely hard case of the sort we are considering. Ethical cases hard to judge in sense (ii) are hard because we find the force of certain compelling arguments pulling us in one direction, and the force of our (honest, enduring, un-self-deluded, critically assessed, etc.) felt responses in the other. And although not always easy to tell, there is often more to the latter's own ethical significance, than their potential for articulation into a general principle. We just need to know where to look to find it.

Now probably the main objection to the view of intuition and sensibility that I am proposing will go something like this:

Of course we may choose to stick with our gut instincts or intuitions. We may say, 'I see there are excellent arguments in favour of X, but I just can't bring myself to believe X is true'. But when we make statements like this we should admit that our rejection of X is not justified; rather, we simply continue to believe, without sufficient justification, that X is false, and so we continue to search for a solid justification for that belief. This is the most we could plausibly say about the intuition that killing innocent threats is permissible, in the light of the apparent concession that philosophers like Thomson have failed to find a satisfying general principle which explains this intuition.¹⁸

¹⁷Good and thoughtful people once supported horrors, like slavery—one of many disturbing examples. But did such good and thoughtful people consider the slaves to be innocent human beings, and moreover, hold their beliefs about slavery in the face of sincere self-examination and careful scrutiny, and together with a sincere and deep commitment to the stringent rights against harm, enjoyed by innocent human beings? If not, then how do they call into doubt the sort of intuitions about Innocent Threats we are considering?

¹⁸I owe this formulation of the objection to a referee for this journal.

7.2 A Strawsonian Reply Based in a Kantian Tradition

But this takes a general position against the relative authority of sensibility that is more controversial than it might first appear to be (especially in ethics). It assumes that our felt or otherwise intuited responses cannot have a decisive authority of their own, backed by, for example, fundamental elements or structural features of our sensibility. It assumes, likewise, that they could never have decisive authority in settling any questions of moral permissibility or parity, not even in cases like the ones at issue.

When the objection is put in the very general way that I have put it above, it opposes both Humean and Kantian traditions in philosophy. It opposes a Humean sensibility-driven reconception of (for example) our inductive inferences: inferences which—even if shown to be baseless by certain excellent arguments—are redeemed by a naturalistic, sentiments-based account of their legitimacy.¹⁹ It opposes Kant's philosophy, which, for us, is the better example because, unlike Hume, Kant gives authority and power to both pure reason and sensibility, insisting only on the interdependence of the two. The first *Critique* anchors the forms of reasoning (other than the purely logical forms) in the forms of our sensibility.²⁰ This has the result of making the conclusions of such otherwise compelling forms of reasoning as Leibniz's doctrine of the identity of indiscernibles²¹ unacceptable because they are arrived at without regard for these independent sensible constraints on reasoning.²² More relevant to our concerns here, the second *Critique* makes feeling constitutive of the practical effect, and thus ethical relevance, of pure reason on human beings.²³ And, to come up to date, this philosophical tradition has had its most able recent exponent in P. F. Strawson, who argues that, as a whole, our ethical sentiments (e.g. resentment and gratitude) and their related judgments could never be challenged or affected by any purely rational line of discursive argumentation, since their authority and normative force derive from certain fundamental and pervasive human feelings.²⁴

This is a richly varied and complicated tradition that I'm alluding to here, but it has a common strand: our sensibility, including our feelings and emotions, can put legitimate, decisive constraints on our discursive reasoning, no matter how compelling and impeccable the reasoning might

¹⁹See David Hume's *Enquiry Concerning Human Understanding*, §V [Hume 1975 (1777)]. Cf. Strawson [1974: 23n].

²⁰See the division entitled Transcendental Analytic in the *Critique of Pure Reason*, and the especially helpful appendix there: On the Amphiboly of Concepts, etc. [Kant 1998 (1781/1787)].

²¹See 'On the Amphiboly of Concepts' [ibid.].

²²As he tries to illustrate most famously in the chapter of the *Critique of Pure Reason* entitled The Antinomy of Pure Reason [ibid.]

²³See chapter 3 ('On the Incentives of Pure Practical Reason') of the *Critique of Practical Reason* [Kant 1996 (1788)], a chapter I discuss at length in Guevara [2000].

²⁴According to Strawson, it is psychologically impossible for us to bring this aspect of human sensibility up for rational assessment as a whole, or else, even if we *could* bring it up for assessment, no purely rational assessment would decide its fate [1974: 13n, 23, and 23n]. So, e.g., we can question whether your resentment and related judgments are appropriate given that they seem directed towards someone who has acted out of blameless ignorance and good will, but not whether resentment and related moral sentiments are justifiable in general. Thus no argument (however compelling) that would call into question their legitimacy in general (e.g. an argument showing that we are totally determined in everything we do by causes extending far out of our control) could in fact do so.

seem to be without those constraints. On this view, there might very well be, then, a compelling argument for some ethical claim X, but one that ignores the independent constraints of sensibility, including those imposed by our felt responses. In such a case (the view is that) X might be legitimately rejected for being inconsistent with these constraints of sensibility.

Perhaps then my view of intuition can find in this philosophical camp shelter enough to shield it for the time being from the very general objection above. This leaves work to be done, of course. I will concentrate on what is manageable in the space that remains.

The particular felt constraints that I am trying to defend are not like the Humean sceptical constraints on reason. I am not promoting sensibility on the basis of a general scepticism about the authority and efficacy of reason. I am proposing only a check on discursive reason in some hard cases. The check that I am proposing has its most direct and evident affinities with Strawson,²⁵ who argues for the irrelevance of any form of reasoning, however compelling and otherwise legitimate, that would establish conclusions contrary to the ethical judgments arising from certain inescapable and fundamental features of human sensibility. I echo something like this Strawsonian idea in my own view when I suggest that what informs our intuitive responses to certain hard cases are feelings that make it impossible or profoundly difficult (e.g. inhumane or deeply alienating) to live by judgments contrary to those feelings. But unlike Strawson (and Hume) I do not adduce feelings that are plausibly universal, original, inescapable for all humans (much less part of the transcendental conditions of sensibility, or something along grand Kantian lines). I have, after all, been engaged with authors who seem not to share them exactly, and in any case, know of many that do not share them.

The feelings that might lead one to become, for example, a strict pacifist, serve me better than feelings, like resentment, that all human beings feel from time to time. I am arguing for the violent option in Innocent Threats, but set aside for the moment that one obvious disanalogy in the analogy I'm about to draw with pacifism.

Whatever other reasons exist for being a pacifist, I believe that a basic and non-negotiable repugnance to harming, especially killing, is often at the root of strict pacifism. The feelings here run at least as deep as any repugnance to treating Innocent Threats and bystanders on a par. I take it, too, that some pacifists of this sort have subjected their position to stringent critical scrutiny and self-examination. That their pacifism has been tested by time and circumstance, and that it is part of a way of life otherwise admirable even to those who themselves feel a repugnance for pacifism or reject it on principle or both. It's not, of course, that confirmed pacifists can never be tempted to violence or indulge their pacifism from dubious motives or whatever, but that their feelings run deep, and may be inescapable for them. They may be the sort of feelings that make one's life meaningless or

²⁵Two referees for this journal suggested the comparison to Strawson. Both also suggested a comparison to John McDowell, especially his *Virtue and Reason* [McDowell 1998: chapter 3]. I believe there are productive and interesting alliances for me to make there too. But they are too complicated to discuss here.

intolerable when not honoured, however compelling the arguments to violate them might be in certain cases.

This sort of sensibility, a sort that many of us have about the right to self-defence against Innocent Threats and (especially) Aggressors, can place legitimate and decisive constraints on a person's way of life and ethical judgment. Moreover, I think this offers a more satisfying and interesting way of comprehending the meaning of our intuitions on this issue: i.e. as bottom-line, non-negotiable feelings, tested and examined, and part of an, in general, uncontroversial and informed position on killing and harm, etc.

This indicates how I would proceed more specifically against the general objection quoted above, and what I mean to evoke with questions about what it would be like to live out the idea that it is *impermissible* to lethally defend oneself against Threats and Aggressors, because on a par with killing mere bystanders. Our sensibility might simply not allow us to take seriously any argument that we must treat bystanders and Threats/Aggressors as morally on a par. When it comes to certain feelings like these, they might be no less inescapable and non-negotiable than the sort adduced by Strawson (or Hume). To try to escape, engineer or negotiate them away is to strike at the core of one's sense of value, including the sense of one's own personal worth.

7.3 *A Response to Otsuka's position*

This might also help indicate the sort of response I would develop against another natural objection: namely, what about those good and thoughtful people who are willing to follow the argument to the end, whatever their intuitive sense? And also about those who have a different intuitive sense, from the one I am defending?

I'll go out on my armchair limb and say that I don't think anyone has the intuitive sense that the two cases—Innocent Threat and bystander—are morally on a par, especially not if we include Innocent Aggressors among the Innocent Threats. This does not imply that everyone's sense is that we can kill the Threat in self-defence: our strict pacifist might have the sense that the two are not on a par (for all their similarities), and yet feel a deep moral repugnance to killing the Threat (because of her general repugnance to killing anyone, even culpable threats).

A few people, such as Otsuka, may, whatever their sense, be willing to follow the (admittedly forceful) line of argument out to the end and accept that the two types of killing—killing of innocent bystanders and killing of Innocent Threats—are on a par morally. They may even be convinced, as Otsuka is [1994: 75, 89], that the Innocent Aggressor is on a moral par with the mere bystander too. For although Innocent Aggressors, unlike Innocent Threats, are trying or intending to kill you, they are hardly more responsible for that fact than the other is for being a threat to you (or for his body being one, or whatever). But I think *everyone*, at some point, feels the strain that this line of argument puts on sensibility, even when the argument seems compelling to them. The tenor of McMahan's especially comprehensive

essay [McMahan 1994]—which draws conclusions very similar to Otsuka’s—is a kind of evidence for this claim. And even Otsuka shows signs of the strain of his argument (in speaking of being prepared to ‘swallow’ certain particularly unpalatable conclusions) [1994: 89]. An interesting question to pursue, interlocutor by interlocutor, and in a variety of ways, depending on the particular interlocutor of this type, is the question of the full meaning of the strain this position creates for almost everyone. For example, does it show that one could not, at some point, really live or imagine living with all the consequences of the position, no matter how good the arguments?

However, I cannot dismiss the beliefs in question just because those who hold them are themselves uncomfortable with some of the consequences of having to live by them, or because they themselves admit it is difficult to accept moral parity between bystanders and Threats. Perhaps legitimate differences in sensibility force on me a kind of relativism in the ethics of these hard cases: a relativism that makes sense of the value of both attitudes (and ways of life), different as they are. The analogy with pacifism might help here. But so much for sketches of future work.

8. Conclusion

I have described and partly defended a position on Innocent Threats that no one in the literature I have been discussing calls attention to or takes seriously enough when they do. I have tried to communicate the appeal of this alternative, and this especially: that we must not think of intuition as always a mere gut response, and thus as at best a mere and fragile starting point for reasoned discovery of principles that are to be followed, wherever they lead; rather, in ethics, at least, we must have equal concern for how much, how radically, those principles may alienate us from what seems right in specific cases. In hard cases, for sure, we must take seriously the ethical significance of a crucial exercise in imagination, such as imagining what it would really be like for us to try to live by the idea that the killing of Innocent Threats or (worse) Aggressors is on a moral par with the killing of a mere bystander. These considerations should not, at any rate, simply be ignored or downplayed by those who are willing to follow the argument wherever it leads.

Now McMahan, who takes a position very similar to Otsuka’s about the impermissibility of killing Innocent Threats and Aggressors (all of which he infelicitously dubs IAs (Innocent Attackers)), at least struggles with the question of ‘... what motivates us to believe so strongly in the permissibility of self-defensive killing of an IA even though we can provide no fully persuasive defense of our belief’, as he says [1994: 288].

He tries out various explanations. In the end he focuses on one, namely that our belief is based on our recognition of the significant social utility there is in treating the self-defensive killings of Innocent Threats and Innocent Aggressors as permissible (and thus more like the paradigmatic killing in self-defence against a culpable aggressor than the killing of mere

bystanders in self-preservation). We all recognize how difficult it would be in practice to be certain of the innocence of an Innocent Threat or Aggressor in the relevant cases, for example. And because the cases are so rare and similar, there would be great social utility in internalizing a rule that permitted killing in these cases.

There are various reasons for not being satisfied with this explanation of our intuition that killing in the Balcony Case, etc., is permissible. The main reason McMahan is unsatisfied with it is that it strongly suggests that the explanation of the wrongness of killing *mere bystanders* is *also* based on a convention we adopt only for the sake of social utility. After all, we have not been able to discover any intrinsic feature of the cases, or general theoretical principle covering the cases, that would distinguish the two types of killing morally. So, McMahan believes that the more reasonable alternative is to accept that the common intuition about the Balcony Case is in fact deontological, or non-consequentialist, in character, and therefore entirely mistaken or deeply confused. Now, of course, I do not believe it is mistaken or confused. I have urged (with the Kantian undertones now made explicit) that our judgment about the cases *not* be determined automatically by the seemingly inevitable implications of closely reasoned analysis of the concepts, *not* when the reasoning is so difficult to square with the sensibility and imagination of those who might very well be showing good judgment in resisting those implications.

This view ought to be taken especially seriously by anyone who is holding out, against the obstacles presented by the casuistry, for an articulate and principled distinction between Innocent Threats (and the like) and mere bystanders—one that overcomes on general principle the bystander-like faultlessness of the Threats. For the role of intuition is obviously quite serious for such holdouts: intuition functions for them like an authoritative signpost indicating the (difficult) way to a philosophical principle. But I don't see why they should believe that intuition is necessarily any less serious, i.e. authoritative, if their hunch that it will lead to a principle is wrong.²⁶

University of California, Santa Cruz

Received: February 2009

Revised: November 2009

REFERENCES

- Daniels, Norman 2008. Reflective Equilibrium, *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), ed. Edward N. Zalta, URL= <<http://plato.stanford.edu/archives/fall2008/entries/reflective-equilibrium/>>.
- Guevara, Daniel 2000. *Kant's Theory of Moral Motivation*, Boulder: Westview Press.

²⁶I would like to acknowledge the help I received on this paper from Julie Tannenbaum, Jonathan Ellis, Judith Baker, and all the referees for this journal who commented on it. Tannenbaum, like the referees, was especially helpful in getting me to understand Judith Thomson's view of rights, and in contributing to many improvements in the style and presentation of the paper. My colleague, Jonathan Ellis, walked me through painstakingly detailed comments on a late draft of the paper, which improved the clarity of my presentation, and contributed to the strengthening of the argument at crucial places. Judith Baker provided very insightful comments on an earlier version, and was the first to make me feel confident about my general line, which I gratefully acknowledge.

- Hume, David 1975 (1777). *Enquiry Concerning Human Understanding*, in *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. L. A. Selby-Bigge and P. H. Nidditch, Oxford and New York: Oxford University Press.
- Kamm, F. M. 1992. *Creation and Abortion*, New York: Oxford University Press.
- Kamm, F. M. 2007. *Intricate Ethics*, New York: Oxford University Press.
- Kant, Immanuel 1998 (1781/1787). *Critique of Pure Reason*, ed. Paul Guyer and Allen Wood, Cambridge and New York: Cambridge University Press.
- Kant, Immanuel 1996 (1788). *Critique of Practical Reason*, ed. Mary J. Gregor, Cambridge and New York: Cambridge University Press.
- Knobe, Joshua and Shaun Nichols, eds, 2008. *Experimental Philosophy*, New York: Oxford University Press.
- Ludwig, Kirk 2007. The Epistemology of Thought Experiments: First Person versus Third Person Approaches, *Midwest Studies in Philosophy* 31: 128–59.
- McDowell, John 1998. *Mind, Value and Reality*, Cambridge: Harvard University Press.
- McIntyre, Allison 2001. Doing Away with Double Effect, *Ethics* 111/2: 219–55.
- McMahan, Jeff 1994. Self-Defense and the Problem of the Innocent Attacker, *Ethics* 104/2: 252–90.
- Nozick, Robert 1974. *Anarchy, State and Utopia*, New York: Basic Books, Inc.
- Otsuka, Michael 1994. Killing the Innocent in Self-Defense, *Philosophy and Public Affairs* 23/1: 74–94.
- Quinn, Warren 1989. Actions, Intentions, and Consequences: The Doctrine of Double Effect, *Philosophy and Public Affairs* 18/4: 334–51.
- Strawson, P. F. 1974. Freedom and Resentment, in *Freedom and Resentment and Other Essays*, London: Methuen and Co. Ltd: 1–25. Originally published in *Proceedings of the British Academy* 48, 1962.
- Thomson, Judith Jarvis 1971. A Defense of Abortion, *Philosophy and Public Affairs* 1/1: 47–56.
- Thomson, Judith Jarvis 1991. Self-Defense, *Philosophy and Public Affairs* 20/4: 283–310.