

# **Ignorance, Beneficence, and Rights**

Kieran Setiya

MIT

September 29, 2018

(Draft; do not cite without permission)

The Trolley Problem is the problem of explaining the moral contrast between two hypothetical cases.<sup>1</sup>

In Bystander, a runaway trolley is heading down a track towards five strangers, who will be killed if you do nothing. The trolley is approaching a fork. If it is redirected by the push of a button, it will swerve on to a side track, where it will kill one stranger instead of five. The circumstance is otherwise unexceptional. You have no special obligation to the strangers. They are not responsible for the runaway trolley and their life or death would have no unusual consequences, good or ill. You are a bystander at the button. What should you do?

Footbridge is similar, except that the trolley can be stopped only by pushing a button that drops someone through a trapdoor on a footbridge over the track, where they will be killed by the trolley but halt its progress towards the five. Again, you are bystander at the button. What should you do?

Many believe that it is permissible to push the button in Bystander, but not in Footbridge. Why? In each case, you must choose between killing one stranger and letting five others die. Does it matter that in Bystander you redirect an existing threat? Or that, in Footbridge, the death of the one is a means by which the five are saved? An increasingly baroque array of cases can be set against these claims and some despair of any resolution.

---

<sup>1</sup> The problem is due to Thomson 1976, drawing on Foot 1967.

Perhaps we should question, instead, what the Trolley Problem presumes: that there is a moral difference between Bystander and Footbridge. In her most recent discussion of the problem, Judith Thomson (2008) concludes that it is wrong to push the button in Bystander, just as it is in Footbridge. Others conclude that it is permissible to push the button in Footbridge, despite the consensus to the contrary.

My sympathies lie with Thomson, though I will not defend them here. Instead, I will focus on our reasons not to push the button in Footbridge. I will argue that they are distinctively relational: they involve the individuals they concern in a particular, puzzling way. The puzzle has been neglected, but it is one we must solve in explaining the nature of rights.

The argument turns on a cousin of Footbridge recently explored by Caspar Hare (2016). In what I will call the Opaque Footbridge case, you know the people involved but do not know which locations they are in. Hare argues that you should push the button in Opaque Footbridge. His aim is not to establish that you should push the button in the original Footbridge case or that it is permissible to do so.<sup>2</sup> It is to shed light on the relationship between morality and reasonable beneficence. The title of his essay is “Should We Wish Well to All?” His answer to the question is yes.

Hare’s argument is illuminating, but I do not accept the conclusions he draws from it. In what follows, I argue, first, that there is no way to quarantine his verdict in Opaque Footbridge from the permissibility of pushing the button in Footbridge, which few concede. As I explain in section 2, Hare’s argument is more radical than he believes. It has implications not just for what we should do in conditions of ignorance but for the standard Footbridge case.

The relation between these verdicts sheds light on the nature of practical reason. In his appeal to ignorance or opacity, Hare exploits the difference between knowing the existence of reasons for action and knowing what the reasons are. Elsewhere, I have explored the following claim:

---

<sup>2</sup> For reticence about Footbridge, see Hare 2016: 466-7, and section 2 below.

SPECIFICITY: It can be rational to respond more strongly to a fact that is a reason with a certain weight than to the fact that there is a reason with that weight. (Setiya 2016: 4)

Though I believe this claim is true, Hare overestimates its scope and thus the scope for conflicting verdicts in Footbridge and its Opaque variant.

Must we conclude that it is permissible to push the button in Footbridge? I say no. As I explain in section 3, we can resist Hare's verdict in Opaque Footbridge by divorcing rights from interests. Section 4 returns to Specificity, asking what we learn from Opaque Footbridge about the reasons not to push the button in the original Footbridge case. The contrast between these situations shows the particularity of rights: the way in which they focus on particular individuals. Access to the reasons rights involve depends on personal acquaintance. Yet the nature of such acquaintance remains obscure.

Before I turn to these claims, I describe Hare's argument in more depth.

### **1. Opaque Footbridge**

In Opaque Footbridge, six people you know are caught up in a trolley case, five on the track, one on the bridge. They know their locations but you don't, and you cannot find out which of them is where.<sup>3</sup> As before, there is a button that will open a trapdoor on the footbridge, dropping whoever is up there in front of the trolley, where they will be killed by its impact but halt its progress towards the five. Hare argues that concern for each of those involved counts in favour of pushing the button. If we give them alphabetical names from A to F, we can see that, by your lights,

---

<sup>3</sup> This case is adapted from Footbridge with Suitcases (Hare 2016: 454-5), with the adjustment that the others know where they are. According to Hare (2016: 466), this detail makes no difference. Hare's argument is a relative of Thomson's (1990: Ch. 7) appeal to hypothetical advance consent, though she was interested in risk, not uncertainty, and in Bystander, not Footbridge. We will come back to consent in section 3.

pushing the button will improve A's prospects from a 5/6 chance of death to just 1/6. It is true that pushing the button will change the potential cause of death, from being hit by a runaway trolley to falling from a bridge in front of a trolley as a result of your intervention. It also has the potential effect of using A's death as a means to save more lives. But from the standpoint of concern for A, these factors seem irrelevant. Why should it be better for A to die on the tracks than falling from a bridge to save the five? The upshot is that, in Opaque Footbridge, concern for A alone should lead you to push the button. The same is true of concern for B, and C, and all the rest. There is no need to weigh conflicting interests. Beneficence speaks with a single voice.

This marks a contrast with the original Footbridge case (Hare 2016: 466). Again, suppose you know the six involved, from A to F. If you know that F is on the bridge, concern for each is not unanimous. There is no way to argue that you ought to push the button without comparing or combining claims. Beneficent concern is simply divided. Concern for F speaks against pushing the button, concern for the others speaks in favour. This conflict cannot be ignored.

It remains true in Opaque Footbridge that pushing button A might save A's life by killing someone else. But we can abstract from this, as Hare explains. He gives a further argument for his conclusion, that you should push the button in Opaque Footbridge, by way of a more intricate scenario, in which each button affects only the person it names. In Two Parallel Tracks, Six Buttons, and All Know Their Place, two tracks run in parallel (Hare 2016: 466). On each track is a trolley, heading towards five boxes. Above each track is a footbridge on which there is a single box. A to F are in the boxes in the vicinity of the first track. They know their locations but you don't, and you cannot find out which of them is where. The boxes in the vicinity of the second track are empty, except for the one on the bridge, which is full of sand. You cannot stop the trolley on the first track, which will run through the boxes in front of it, killing anyone inside. But you can push any combination of six buttons, marked "A" to "F." If you press "A," you move A's box from its position on the first track to the corresponding position on the second track, without learning where that is. Likewise for the other buttons. If you push any button, a trapdoor opens on the second bridge, dropping the box in that location on to the track, halting the second trolley.

If someone is in the box when this happens, they will die. If you do not stop the second trolley, it will run through the boxes in front of it, killing anyone inside. Hare makes the following argument about this case.<sup>4</sup>

(1) Regardless of whether you push the other buttons, you should push “A”; and regardless of whether you push the other buttons, you should push “B”; etc.

(2) For any composite action,  $A_1 \dots A_n$ , if you should do  $A_1$  regardless of whether you do  $A_2$  through  $A_n$ , and ... and you should do  $A_n$  regardless of whether you do  $A_1$  through  $A_{n-1}$ , then you should do  $A_1 \dots A_n$ .

(3) You should push all six buttons.

Presumably, if you should push all six buttons, you should push them even if the only way to do so is to push them all at once. It makes no difference if the buttons are linked together, or if instead of six buttons, there is only one, which does the work of six (Hare 2016: 465). In that case, you should push the single button, dropping an unknown victim in front of a trolley in order to save the other five. But if you should push the single button in this variant of Two Tracks, you should push the button in Opaque Footbridge. The effects of doing so are morally just the same. Hence:

(4) If you should push all six buttons in Two Tracks, you should push the button in Opaque Footbridge.

This ingenious argument will occupy us for some time. We won't be concerned with premise (2) or premise (4). But why accept premise (1)? Hare's thought is that, regardless of what

---

<sup>4</sup> What follows is not verbatim; it simplifies Hare 2016: §4.

else you do, pushing “A” is in A’s interests, by your lights. It improves her prospects from a 5/6 chance of death to just 1/6. As in Opaque Footbridge, pushing the button will change the potential cause of death, from being hit by a runaway trolley to falling from a bridge as a result of your intervention. If other buttons have been pressed, it may also change the effects: if A is in the box on the bridge and you push “A,” he will end up blocking the trolley on the second track, preventing it from killing anyone else. Again, however, these factors seem irrelevant to concern for A. What is more, while the button in Opaque Footbridge might save A by killing someone else, pushing button “A” will not affect anyone but A. There is no risk that, in pushing it, you will cause someone else to die who would otherwise live, or save A’s life by killing them. Since only A is affected, it is only concern for A that should count: the only relevant affect of pushing the button is to improve A’s prospects of staying alive. The same is true of pushing button “B” in relation to B’s prospects, and of button “C” in relation to C, etc. Hence, premise (1) is true, and the argument is sound.<sup>5</sup>

We will return to this premise below. First, we need to ask how much the argument shows. Does it follow from Hare’s conclusion, that you should push the button in Opaque Footbridge, that you should push the button in Footbridge or that it is permissible to do so? In section 2, I argue that it does.

## **2. Three Grades of Specificity**

As Hare points out, the argument above has no direct application to Footbridge, since it depends on ignorance of who will die if the button is pushed (Hare 2016: 466). In Footbridge, as we will

---

<sup>5</sup> It is true that, as you press more buttons, the cumulative probability that you have killed someone goes up. But it does not make sense to care about this. Imagine being in a situation in which you should risk someone’s life in order to produce some positive effect. If the situation repeats itself, you should act in the same way, no matter how many times it happens. It would be absurd to stop at a certain point on the ground that, while the marginal risk of killing someone is the same as always, the cumulative risk of having done so is now too high.

imagine it, you know that F is on the bridge, so that concern for F conflicts with concern for the interests of A to E. Hare suggests that his argument leaves open what to say about this case. Perhaps you should push the button in Footbridge. Perhaps it is permissible to do so but permissible to refrain; this is Hare's own view (Hare 2016: 467). Or perhaps it is wrong to push the button, as orthodoxy holds.

I do not think we can be so flexible. Hare's verdict in Opaque Footbridge may not strictly entail that it is permissible to push the button in Footbridge; but it is inconsistent with knowing that you shouldn't.

Suppose you know that you should not push the button in Footbridge. (Here and elsewhere, "should" means "rationally should": this is what it would be rational to do, given your beliefs. We stipulate that the beliefs are true and that they include the description of the case.) In Footbridge, you have decisive reason not to push. What should you want yourself to do in Opaque Footbridge? From the point of view of full information, Opaque Footbridge looks the same as Footbridge. The reasons that lead you to refrain from pushing the button in Footbridge should make you prefer that you refrain in Opaque Footbridge. That is what a fully informed version of you, knowing who is on the bridge, would want you to do. Thus, in Opaque Footbridge:

(a) You know that, with full information about your circumstance, you should prefer that you refrain from pushing the button.

But now consider:

(b) If you know that, with full information about your circumstance, you should prefer X to Y, it is not the case that you should prefer Y.

It follows from (a) and (b) that it is not the case that you should prefer that you push the button in Opaque Footbridge. And so it is not the case that you should do so. That conflicts with the

argument of section 1. If we are forced to accept that argument, consistency requires us to relinquish (a). We must deny that, in Footbridge, you should not push the button: it is permissible to do so.

The only way out, the only way to drive a wedge between the permissibility of pushing the button in Footbridge and the claim that you should do so in its Opaque variant, is to cast doubt on (b). That would be Hare's response. Addressing a similar principle cited by Marc Fleurbaey and Alex Voorhoeve (2013: 117-21), he claims that "where things of incommensurable value (like people's lives) are at stake, it makes perfect sense to resist the judgments of your better informed self" (Hare 2016: 456-7). He cites in support of this the "prospectist" decision theory he has argued for elsewhere (Hare 2013: Ch. 3). But he overestimates the scope for divergence between the rational response to knowledge of the weight of reasons and the rational response to reasons themselves.

Consider again the claim, suggested above, that such divergence is possible:

**SPECIFICITY:** It can be rational to respond more strongly to a fact that is a reason with a certain weight than to the fact that there is a reason with that weight.

When the fact that  $p$  is a reason for you to  $\phi$ , Specificity claims, it can be rational to respond more strongly to knowledge that  $p$ , knowledge that specifies the reason, than to knowledge that there is a reason of this kind. In the first case, you are in a position to reason from the fact that  $p$ , to be moved by the reason itself; in the second case, your relation to the reason is mediated or indirect. This opens space for divergence between the attitudes you should have, given what you know, and the attitudes you know you should have with full information. It opens space for exceptions to (b).

The question we need to ask is how much space there is. Specificity takes different forms. In relation to preference, we can distinguish three. In each case, we compare the preference you should have in your actual situation with the preference you should have about that situation from the perspective of full information. What should you want for your less-informed self?



MILD SPECIFICITY: You should prefer X to Y despite knowing that, with full information about your circumstance, it would be rational to prefer Y.

MODERATE SPECIFICITY: It is rational to prefer X to Y despite knowing that, with full information about your circumstance, you should prefer Y.

EXTREME SPECIFICITY: You should prefer X to Y despite knowing that, with full information about your circumstance, you should prefer Y.

The names for these phenomena are tendentious: it is not obvious from their structure how Mild and Moderate Specificity should be ranked. But the claims are increasingly peculiar. In brief: Mild Specificity is possible, Moderate Specificity is contentious, and Extreme Specificity is absurd. But it is only Extreme Specificity that falsifies (b), which states that it is impossible.

Mild Specificity is involved in the phenomenon “opaque sweetening,” which motivates Hare’s prospectist decision theory (Hare 2013: Ch. 3) Suppose you will receive tickets to a concert or tickets to the theatre. The play looks wonderful, the musician is brilliant – you expect that you’d enjoy them both – and the ticket price is the same. It would be rational to prefer either one. What is more, it would be rational to prefer the concert tickets at the cost of an extra dollar, or to prefer the theatre tickets when they cost a dollar more. In Ruth Chang’s terminology, the relation between your reasons – the fact that this ticket will admit you to the play; or that it will admit you to the concert – is one of normative parity: not equally strong, but also not outweighed (Chang 1997). Now imagine a variant of this case in which you will receive one of two opaque boxes, with concert tickets in one, tickets to the theatre in the other. You ought to be indifferent. But if the box on the left comes with an extra dollar, you should prefer to get that box. So says Hare, and I agree. The cost is Mild Specificity. You should prefer the box on the left with the extra dollar despite

knowing that, with full information about your circumstance – knowing what is in each box and thus the reason to want it – it would be rational to prefer the box on the right.

Moderate Specificity goes beyond this. Here you know that, given full information about your circumstance, you should prefer Y. In terms of reasons for preference: you know that there are facts about your circumstance that provide decisive reason to prefer Y; if you knew these facts, it would be irrational to prefer X. If you know all this, and still prefer X, you are guilty of *akrasia*, preferring what you know there is decisive reason against. This is not the kind of allegedly rational *akrasia* in which you prefer X to Y despite the false belief that you should prefer Y.<sup>6</sup> It is a case in which you know that, with full information about your circumstance, you should prefer Y. And yet you do not prefer it.

The existence of Moderate Specificity cannot be motivated by examples of opaque sweetening. Is it possible? I think so.<sup>7</sup> In section 4, I will argue that Moderate Specificity is involved in the best account of Footbridge and its Opaque variant. There is a distinction between knowing the existential fact that there are reasons to want something and knowing what those reasons are. The idea behind Moderate Specificity is that it is rational to be moved more strongly by knowledge of the second, more specific kind, than knowledge of the first. When you know that there is decisive reason to prefer X to Y but you do not know what the reasons are, it may be rational to prefer Y, giving more weight to the admittedly weaker reasons of which you are aware.

This is undoubtedly controversial. Many will assume that the kind of *akrasia* involved in Moderate Specificity is irrational. Yet however bad things look for Moderate Specificity, Extreme Specificity is worse. In Extreme Specificity, you know that, given full information about your circumstance, you should prefer Y. You know that there are facts about your circumstance that provide decisive reason to prefer Y; if you knew these facts, it would be irrational to prefer X. And

---

<sup>6</sup> See McIntyre 1990, Arpaly 2000.

<sup>7</sup> I have argued elsewhere that we must accept Moderate Specificity in order to make sense of the affirmation of one's actual life in the face of better alternatives (Setiya 2016: §II).

yet it is not only rational to prefer X but irrational not to. *Akrasia* is rationally required! This strikes me as incredible.<sup>8</sup> If you know that, with full information about your circumstance, you should prefer X to Y, it is not the case that you should prefer Y. And that is premise (b).<sup>9</sup>

Hare is right to allow for divergence between the preferences you should have, given what you know, and those you know you should have with full information. He is right to deny that, in general, your preferences must align with those of your better-informed self. But he is wrong to think that this leaves room for the orthodox view of Footbridge. The degree of Specificity required to reconcile Hare's verdict in Opaque Footbridge with the claim that you should not push the button in Footbridge is untenable. If he is right about Opaque Footbridge, it is permissible to push the button in the Footbridge case. This is the view he favours anyway. If I am right, the argument in section 1 commits him to it.

While Hare may be happy with this commitment, others will demur. Like me, they may feel that his conclusion about Opaque Footbridge is defensible only if it can be insulated from the permissibility of pushing the button in Footbridge; and this cannot be done. What we learn from (a) and (b) is not that we were wrong about Footbridge, but that we must reject Hare's argument about its Opaque variant. We must revisit premise (1).

---

<sup>8</sup> Suppose you can bet on the winner of a race at very poor odds. You should prefer not to take the bet. But if you had full information, including information about who will win the race, you should prefer to take it. Is this a case of Extreme Specificity? No. The question is not what you should prefer to do if you had full information, but what you should prefer to happen in your actual circumstance. Even with full information, you should not want yourself to take the bet in a condition of ignorance, since you would be forced to guess about the winner. It is true that, if you knew that Mr. Ed will win, you should want yourself to bet on Mr. Ed. But in your condition of ignorance, you are not in a position to know that with full information, you should want yourself to bet on Mr. Ed. So again, we have no instance of Extreme Specificity.

<sup>9</sup> In Newcomb's Problem, (b) conflicts with the view that you should prefer to take one box, given that, if you knew what was in the boxes, you should prefer to take both. I regard this not as an argument for Extreme Specificity, but as an objection to one-boxing.

### 3. Rights and Interests

Hare subsumes the argument for premise (1) under a principle of “defeasible privacy”:

When you know that an action will affect the well-being of only one person, and you know who that person is, then it is morally appropriate, in the absence of powerful defeating considerations, to act out of reasonable concern for that person. (Hare 2016: 463)

He does not specify everything that would count as a powerful defeating consideration, but he insists that considerations about cause of death are not sufficient. If you push button “A,” there is a chance that A will be killed by your intervention; if you refrain, she is likely to die without it. But this is not enough to outweigh the force of concern for A’s well-being. Thus:

(1) Regardless of whether you push the other buttons, you should push “A”; and regardless of whether you push the other buttons, you should push “B”; etc.

What defeasible privacy brings out is that, unlike Footbridge, Two Parallel Tracks does not confront us with a conflict of interests. Many believe that we have rights against interventions that sacrifice our interests for the benefit of others. We are not just units in the greater good; to treat us that way is to ignore the separateness of persons. When you push the button in Footbridge, you may violate such rights. That would explain why it is wrong to do so. In the scope of defeasible privacy, however, the separateness of persons is not at stake. You know who is affected by your action and there is no question of protecting their interests from the claims of others. What serious objection can be made to pushing “A”?

This is a good question, but I believe it has an answer. Not all rights protect the interests of those who have them. Others protect their autonomy or self-governance, or call upon us to treat them with respect. There are powerful reasons not to intervene in others’ lives even when

they would benefit from intervention. That is why paternalism is often wrong. In many cases, one should not act on reasonable concern for another's interests because doing so would compromise their authority over themselves. You cannot confiscate my belongings, or confine my body, whenever I would use them in self-destructive ways. It is wrong to perform surgery on me against my will, even if I will benefit from it.

The independence of rights from interests is pervasive. Think about the contrast between killing someone and letting someone die. It is often assumed that there is a common moral core to these phenomena, an objection that applies to both. The open question is whether and why there is a further objection to killing that doesn't apply to letting die. But as Matthew Hanser has argued, the background assumption is false (Hanser 1995). The objection to letting someone die is grounded in their interests. If someone would die anyway, moments later, were you to save their life, there is little reason to save them. The strength of the reason not to let someone die typically depends on the size of the benefit that saving them would confer. In Hanser's example, a swimmer is threatened by sharks; you can kill the closest shark; but the next shark will get them instead (Hanser 1995: 180-1). Your reason to intervene is very much weaker than it would be if there were only one shark.

The objection to killing is quite different. In general, it is wrong to kill someone even if they would otherwise be killed a moment later. You are as guilty of murder when you are the first assassin to pull the trigger, the others being seconds behind, as when you are the only assassin (Hanser 1990: 183-4). The strength of the reason not to kill is not proportional to the size of the harm you cause. If there is a case in which death would come as a benefit to me, since life is pure suffering, it is still wrong to kill me without my consent. Some rights protect interests, but the right to life – the right we violate in killing someone against their will – does not.

It is thus a matter of great importance whether people consent to being killed. It is arguably permissible to push the button in Footbridge if the person on the bridge agrees to sacrifice her life for the sake of the others. What about Two Parallel Tracks? Here it is significant that, unlike Hare, I began with a case in which, while you are ignorant of their particular

locations, the people involved know where they are. His initial descriptions are different in that the ignorance goes both ways: “Nobody has any idea who is where” (Hare 2016: 455). This means that there is an argument for presumed consent, for thinking that each of A to F would want you to push the relevant button on their behalf (Hare 2016: 457). In that case, it may be true that you should push the buttons. But if it is consent that does the work, this has no bearing on what to do in Footbridge, where the consent of F is not forthcoming. The more interesting case is the one I have described, in which you are ignorant of who is where, but the others are not. Suppose that A is on the bridge, she knows it, and she is hoping that you won’t push “A.” According to Hare, this makes no difference to what you should do. By your lights, there is a 5/6 chance that A wants you to push “A,” saving her life, and a 1/6 chance that she is hoping you won’t; no-one else will be affected. Reasonable concern for A should lead you to push the button and that is what you should do.

I am not convinced. Being killed without consent may not do more harm to A than being allowed to die. But the moral objection to killing is more profound. It is not about A’s interests but her rights. In considering the button, you must weigh A’s prospects of survival against her right to be free of unwanted fatal intervention. Pushing the button improves her odds of living, on your evidence, but risks killing her against her will. You cannot assume that she would want you to intervene, since she may not. I do not say that it would be wrong to push the button in this case, that a 1/6 chance of violating A’s right not to be killed without consent should count for more than a 5/6 chance of saving her life. But I don’t think the reverse is true. You do not have decisive reason to push “A”, or “B,” or any of the rest.

In defending this verdict, let me address some further arguments in support of premise (1). One is that, in preferring not to push button “A,” not out of concern for A’s interests but concern to avoid killing her, you display an indecently selfish desire to keep your own hands clean.<sup>10</sup> This is doubly misleading. First, there is an ambiguity in “concern for A.” We have been

---

<sup>10</sup> Hare mentions dirty hands in passing at Hare 2016: 457; he elaborates in Hare 2013: Ch. 6.

using this to mean concern for A's prospective well-being. Understood in this way, my reluctance to push the button springs from something other than concern for A. But we could read the phrase inclusively, to incorporate concern for A's autonomy. Understood in that way, it is out of concern for A that I may not want to push the button. Second, there is nothing selfish about my desire. In fact, it has nothing to do with me. This comes out when we ask what I would want someone else to do if they were to face Two Parallel Tracks. Should I want them to push "A"? No more than I want to do so myself. My concern for A's right against unwanted intervention is not self-referential or "agent-relative": it affects what I want to be done to A, whether or not it is done by me.<sup>11</sup>

A different argument draws a comparison between pushing button "A" and operating on a patient who faces a 1/6 chance of death from surgery but a 5/6 chance of dying without it (Hare 2016: 459). You should operate on the patient, even without consent. Isn't this analogous to pushing button "A"? No. The difference is that, in surgery, there is a strong case for presumed consent. Where that is absent, it is rational to refrain. This is clear when you know that the patient would object to invasive surgery. But it is also true when you know that she might. Suppose, for instance, that the surgical risk depends on whether the patient has a genetic condition that affects one in six. Those who have the condition will be killed by the operation; others will be fine. You have no idea whether your patient has the condition or not. But she does. There is a 1/6 chance that the surgery would be fatal and that she would refuse it if she could. Can you operate without consent, if consent is unavailable? As with button "A," I do not say that that it would be wrong to do so. But I deny that you should.

Finally, Hare asks you to imagine "what you would do if [A] were someone you cared about [...] your daughter, or mother, or sister, or old friend." (Hare 2016: 460). "I would push [the button]," he reports, "And though I would feel vexed about it, though I would chew my fingers in horror at the thought that I might be killing her, I would not feel guilty about it. This is just

---

<sup>11</sup> I elaborate on this point in Setiya 2018.

between me and her, and I am acting out of concern for her interests as best I can. That is what I ought to do.” (Hare 2016: 460) I suspect that reactions to these remarks will vary. But I want to emphasize one thing: what we are invited to imagine here adds morally relevant detail to the case of Two Parallel Tracks. It is not just people you know and care about who are facing death, but people with whom you have a meaningful relationship and for whose interests you are specially responsible. Paternalistic intervention is easier to justify when it is the life of a relative or close friend in which you intervene. Relationships attenuate the boundaries between us, the claims of our autonomy. It should be no surprise that the case for pushing the button is correspondingly stronger. But this is not a reliable guide to the strength of reasons when such relationships are absent, as they are in standard versions of Footbridge. Let us keep Two Parallel Tracks and Opaque Footbridge as similar to Footbridge as we can.

The conclusion I draw from these reflections is that Hare’s argument for premise (1) is inconclusive. It is not the case that you should push button “A”; and it is not the case that you should push the single button in Opaque Footbridge. Yet I do not think he is entirely wrong. There are two views we could take of these examples. On the first view, you should not push the button in Opaque Footbridge or the buttons in Two Parallel Tracks. The Footbridge verdict applies. On the second view, it would be rational to push the buttons in Opaque Footbridge and Two Parallel Tracks, and rational not to. In that respect, these cases differ from Footbridge. Hare is right to draw a contrast between them but he is wrong about what it is. I incline to the second view. When I consider pushing “A,” weighing A’s interests against her right to be free of unwanted intervention, factoring in the odds, I think it would be rational to push and not to push. What is more, I think it would be rational to take the same attitude to each button, regardless of which ones you push. I conclude that, in Two Parallel Tracks, it is rational to push all six buttons. It makes no difference if, instead of six buttons, there is only one, which does the work of six. And so it is rational to push the button in Opaque Footbridge.



This leaves us with a puzzle. How can it be rational to push the button in Opaque Footbridge if you should not do so in the original Footbridge case? Is there a way to reconcile these claims?

#### 4. Why Not Push?

If you should prefer not to push the button in Footbridge, but it is rational to do so in Opaque Footbridge, these cases present an instance of

MODERATE SPECIFICITY: It is rational to prefer X to Y despite knowing that, with full information about your circumstance, you should prefer Y.

Moderate Specificity is odd. In Opaque Footbridge, you know that there are facts about your circumstance that provide decisive reason to prefer not to push the button; if you knew these facts it would be irrational to prefer otherwise. And yet it is rational not to prefer that you push the button: an instance of rational *akrasia*. If this is to be intelligible, it must be that you are ignorant of the facts that provide decisive reason in Footbridge. It is rational to be moved more strongly by the reasons you know than by the fact that there is decisive reason to prefer that you not push the button, since you don't know what it is.

This tells us something about the reasons that apply in Footbridge. Perhaps surprisingly, the decisive reason is not that you will kill someone if you push the button, that you will do so without their consent, or that you will use them as a means to save more lives. If these were decisive reasons, you would be equally aware of them in Opaque Footbridge and you should act and desire accordingly. But that is not the case.

What, then, is the decisive reason? What you do not know in Opaque Footbridge is *who* will be killed if you push the button. In the standard Footbridge case, you do. As we imagined this case earlier, you know that F is on the bridge. Alternatively, you can see who is on the bridge and

know that *she* will be killed if you push the button. The decisive reason is a proposition about F, or the person on the bridge, that specifies who she is. You know that, in Opaque Footbridge, there is a truth of this kind about whoever is on the bridge, but you do not know the relevant proposition. That is how Moderate Specificity gets a grip.

The hard question is what is involved in knowing who you will kill. Discussions of “knowing who” in the philosophy of language do not look helpful here. On the minimal view, to know who someone is to know an answer to the question “Who is ...?” The answer can be anything you like.<sup>12</sup> On a more demanding view, to know who someone is to know a contextually relevant answer to the question “Who is ...?” that takes the form of a definite description.<sup>13</sup> Propositions of the form “... is the XYZ” specify who someone is. But this does not amount to progress. In Opaque Footbridge, you know that you will kill the person on the bridge if you push the button. Why isn’t “the person on the bridge” a contextually relevant answer to the question “Who will be killed?”

Because he draws a moral contrast between Footbridge and Opaque Footbridge, Hare faces a similar challenge.<sup>14</sup> For Hare, it is because you know who you will kill in Footbridge that it is permissible not to push the button. He suggests that the sort of “knowing who” that makes a difference is knowing facts about what matters in the lives of those involved, about their friends and families, hobbies and careers (Hare 2016: §6). These are the contextually relevant answers to the question “Who is ...?” What blocks the argument for pushing the button is the plurality of values realized by the diverse activities of A to F: these values are incommensurable and stand in relations of normative parity. But this cannot be right, in Hare’s case or in ours. It would not affect the ethics of Footbridge if the people involved were perfect duplicates of one another, identical sextuplets who lead identical, solitary lives. Nor would it matter if they were people you just met and about whom you know nothing at all. Biographical knowledge is irrelevant.

---

<sup>12</sup> See, for instance, Braun 2006.

<sup>13</sup> This is a drastic simplification of the theory proposed in Boër and Lycan 1986.

<sup>14</sup> It is raised by Harman (2015: 870) in a review of Hare 2013.

What is distinctive of the reason not to push in Footbridge is that it is particular rather than general: it not just that you will kill someone but that you will kill this specific individual. This sort of particularity goes beyond mere definite description, as in “the person on the bridge.” Nor is it a matter of rigid designation. “The person who is actually on the bridge” picks out the same individual in every possible world in which it refers to anyone, but the availability of that description in Opaque Footbridge does not put you in contact with the reason not to push in Footbridge.

Alongside treatments of “knowing who,” there is a history of philosophical reflection on “singular thought,” from Frege to the present. Singular thoughts have been conceived in different ways, but they are meant to involve a form of direct cognitive contact with objects that differs from reference by description. This is a more promising place to turn in understanding what is distinctive of the reason not to kill. It may be the singular proposition that you will kill F that provides the decisive reason of which you are ignorant when you merely believe that you will kill the person who is actually on the bridge.

For Russell (1910-11), acquaintance with particulars is what makes them available for singular thought. Russell’s views about this topic evolved over time and are subject to interpretive dispute, but in his earliest phase, he seems to have believed that we are acquainted only with sense-data, universals, and the self. That idea has not fared well. For those who remain sympathetic to Russell, the paradigm of acquaintance is perceptual contact of the sort that sustains demonstrative reference.<sup>15</sup> This is something you used to have with people you have met, and it is something you have with the person you see on the bridge. So long as you remember the them, you are capable of thinking about them in a distinctively, perceptually generated way.

Unfortunately, the idea of a distinctive category of singular thoughts, supplied by perceptual contact, is controversial.<sup>16</sup> Even if it is right, it leaves us very much in the dark. Why

---

<sup>15</sup> See, for instance, Dickie 2015: Ch. 4.

<sup>16</sup> For a recent critique, see Hawthorne and Manley 2012: Ch. 3.

should perceptual contact, past or present, have moral significance? Why should seeing the person on the bridge, or being able to identify them with someone you met before, make it wrong to push the button, where knowing that you will kill the person on the bridge does not? It is difficult to say.

No similar puzzle is raised by opaque sweetening, as discussed in section 2. The facts of which you are ignorant in that case are easy to state: that the box on the left contains the concert tickets; or that it houses tickets to the theatre. There is no mystery about what is involved in knowledge of these facts or why they matter, no need to invoke acquaintance or singular thought. Specificity applies to reasons of this mundane variety. But its application to Opaque Footbridge turns on reasons of a puzzling, elusive kind. The puzzle is how to formulate these reasons: how to explain what is involved in knowing them and why it has moral weight.

Instead of solving this puzzle, I will give it a name. Let “personal acquaintance” be the cognitive relationship that makes someone available for thoughts of the sort that matter in Footbridge.<sup>17</sup> In Fregean terms, it is the basis of distinctive singular concepts that figure in our reasons, as when the fact that you will kill F is a reason not to push the button that goes beyond the fact that you will kill the person who is on the bridge. For non-Fregeans, personal acquaintance is a guise under which we can think of others: what it is rational for us to want depends on what we think in this distinctive way. If we are to make sense of the contrast between Footbridge and Opaque Footbridge, we need to make sense of personal acquaintance and why it matters in the way it does.

Doing so would shed light on the contrast between “identified” and “statistical” victims sometimes noted in public policy. We are more willing to tolerate merely statistical harms, as when a construction project is virtually certain to cause at least one death, by exposing many people to small risks, than harm to victims identified in advance, as when we know who the

---

<sup>17</sup> Elsewhere, I have used “personal acquaintance” for the minimal cognitive relation that makes sense of love, as in love at first sight, and of beneficent concern (Setiya ms.). I believe that these relations are the same.

project will kill. Can this preference be justified? If we are responding to a moral distinction, what could it be? Suggestions include the fair distribution of risk (better to impose small risks on many than great risks on few) and appeal to person-affecting principles (it matters whether a determinate individual would be made worse off).<sup>18</sup> Neither would explain the contrast between Footbridge and Opaque Footbridge. It invokes a further sense of morally significant identification, one that demands philosophical treatment.

In an essay on “Other People” (Setiya ms.), I take steps in this direction, drawing on the work of Emmanuel Levinas.<sup>19</sup> His ethical account of the face of the other can be read as a theory of personal acquaintance. This idea is central to his most well-known book, *Totality and Infinity* (1961), but it appears much earlier, as in “Freedom and Command,” published in 1953:

The being that expresses itself, that faces me, says *no* to me by this very expression. This *no* is not merely formal, but it is not the *no* of a hostile force or a threat; it is the impossibility of killing him who presents that face; it is the possibility of encountering a being through an interdiction. The face is the fact that a being affects us not in the indicative, but in the imperative, and is thus outside all categories. (Levinas 1953: 21)

Levinas insists on the particularity of our relation to the other in the commandment, “Thou Shalt not Kill”:

I must judge, where before I was to assume responsibilities. Here is the birth of the theoretical; here the concern for justice is born, which is the basis of the theoretical. But it is always starting out from the Face, from the responsibility for the other that justice

---

<sup>18</sup> For the first suggestion, see Daniels 2012, and for discussion of the second, Hare 2012.

<sup>19</sup> I am no expert on Levinas, but I have been inspired by his writings. Michael Morgan’s (2007) *Discovering Levinas* is an invaluable guide; I have also been helped by Perpich 2008.

appears, which calls for judgment and comparison, a comparison of what is in principle incomparable, for every being is unique; every other is unique. (Levinas 1982: 104)<sup>20</sup>

If we are to make sense of the contrast between Footbridge and Opaque Footbridge, we need to make sense of Levinas on the face of the other. We need to explain what personal acquaintance is and why it matters. I won't attempt that here. Instead, I will focus on what we learn from the argument that takes us to this point. What we learn is that the right we violate by pushing the button in Footbridge, and the rights we threaten in Opaque Footbridge, correspond to reasons that relate us to others as particular individuals. The reason in question is not that you will kill someone, or that you will kill the person who is actually on the bridge, but involves the other in a more intimate way. A theory of rights must be in part a theory of this intimate relation, of personal acquaintance or the face, and of its moral weight. This feature of rights is brought into relief by what might otherwise seem a trivial exercise in the ethics of people in boxes on tracks.<sup>21</sup>

---

<sup>20</sup> On the particularity of ethics in Levinas, see Morgan 2007: 61, 79-80.

<sup>21</sup> For comments on earlier versions of this material, I am grateful to: Gregory Antill, Samuel Asarnow, Renée Bolinger, Laura Callahan, Eugene Chislenko, Ryan Doody, Kevin Dorst, Julia Driver, Zoë Johnson-King, John Keller, Brendan de Kenessey, Kathryn Lindeman, Casey O'Callaghan, Doug Portmore, Roy Sorensen, Julia Staffel, and Ralph Wedgwood; to audiences at Temple University, the St. Louis Area Conference on Reasons and Rationality and MIT; and to a reader for this journal. Special thanks to Caspar Hare for inspiration and advice.

## References

- Arpaly, N. 2000. On acting rationally against one's better judgment. *Ethics* 110: 488-513.
- Boër, S. E. and Lycan, W. G. 1986. *Knowing Who*. Cambridge, MA: MIT Press.
- Braun, D. 2006. Now you know who Hong Oak Yun is. *Philosophical Issues* 16: 24-42.
- Chang, R. 1997. Introduction. In R. Chang, ed., *Incommensurability, Incomparability, and Practical Reason*. Cambridge, MA: Harvard University Press, 1997: 1-34.
- Daniels, N. 2012. Reasonable disagreement about identified vs. statistical victims. *Hastings Center Report* 42: 35-45.
- Dickie, I. 2015. *Fixing Reference*. Oxford: Oxford University Press.
- Fleurbaey, M. and Voorhoeve, A. 2013. Decide as you would with full information! In N. Eyal, S. Hurst, O. Norheim, and D. Wikler, eds., *Inequalities in Health: Concepts, Measures, and Ethics*. Oxford: Oxford University Press, 2013: 113-128.
- Foot, P. 1967. The problem of abortion and the doctrine of the double effect. *Oxford Review* 5: 5-15.
- Hanser, M. 1995. What is wrong with killing and letting die? *Philosophy and Public Affairs* 24: 175-201
- Hare, C. 2012. Obligations to merely statistical people. *Journal of Philosophy* 109: 378-390.
- Hare, C. 2013. *The Limits of Kindness*. Oxford: Oxford University Press.
- Hare, C. 2016. Should we wish well to all? *Philosophical Review* 125: 451-472.
- Harman, E. 2015. Review of Caspar Hare, *The Limits of Kindness*. *Ethics* 125: 868-872.
- Hawthorne, J. and Manley, D. 2012. *The Reference Book*. Oxford: Oxford University Press.
- Levinas, E. 1953. Freedom and command. Reprinted in *Collected Philosophical Papers*. Translated by A. Lingis. Pittsburgh, PA: Duquesne University Press, 1998: 15-23.
- Levinas, E. 1961. *Totality and Infinity*. Translated by A. Lingis. Pittsburgh, PA: Duquesne University Press, 1969.
- Levinas, E. 1982. Philosophy, justice, and love. Reprinted in *Entre Nous: On Thinking-of-the-Other*. Translated by M. B. Smith and B. Harshav. New York, NY: Columbia University Press, 1998: 103-121.
- McIntyre, A. 1990. Is akratic action always irrational? In O. Flanagan and A. O. Rorty, eds., *Identity, Character, and Morality*. Cambridge, MA: MIT Press, 1990: 379-400.
- Morgan, M. 2007. *Discovering Levinas*. Cambridge: Cambridge University Press.
- Perpich, D. 2008. *The Ethics of Emmanuel Levinas*. Stanford, CA: Stanford University Press.

Russell, B. 1910-11. Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society* 11: 101-128.

Setiya, K. 2016. Retrospection. *Philosophers' Imprint* 16: 1-15.

Setiya, K. 2018. Must consequentialists kill? *Journal of Philosophy* 115: 92-105.

Setiya, K. ms. Other people.

Thomson, J. J. 1976. Killing, letting die, and the trolley problem. *The Monist* 59: 204-217.

Thomson, J. J. 1990. *The Realm of Rights*. Cambridge, MA: Harvard University Press.

Thomson, J. J. 2008. Turning the trolley. *Philosophy and Public Affairs* 36: 359-374.