

# A Kantian Solution to the Trolley Problem

*Pauline Kleingeld*

## Introduction

The well-known moral dilemmas referred to as ‘trolley cases’ involve a runaway trolley that will kill five people if it continues on its course; their lives can be saved, but, in the paradigm scenarios, only at the cost of the life of another person. The philosophical interest of these cases lies in the fact that most people (including many philosophers) regard it as morally *permissible* or even *required* for a bystander to save the five by diverting the trolley onto a different track where it will kill one person, whereas most regard it as *impermissible* to save the five by pushing a person off a footbridge into the path of the oncoming trolley. Which moral principle, if any, can account for the salient difference between these cases? This question has proved extremely difficult to answer, and this difficulty is known as the ‘Trolley Problem.’ At least in part because a solution to the problem seems impossible to many (though not all),<sup>1</sup> attention has recently been shifting towards psychological explanations

<sup>1</sup> Some have argued for a solution based on the Doctrine of Double Effect. Much of the related debate focuses on whether its core distinction between ‘intended’ and ‘foreseen but not intended’ consequences can distinguish effectively between the Bystander and the Footbridge case. In this chapter, I do not take a stance on whether the Trolley Problem can be solved by the Doctrine of Double Effect, nor do I assess Frances Kamm’s Doctrine of Triple Effect (Kamm 2007). The argument of this chapter does not rule out the possibility of there being more than one solution.

of some of the intuitive judgments involved,<sup>2</sup> and an increasing number of authors now condemn the very use of trolley cases in ethics.<sup>3</sup>

In this chapter, I explore a new avenue for solving the Trolley Problem by focusing on the *practical reasoning* attributed to agents in trolley cases. In particular, I develop a solution to the Trolley Problem in terms of the Kantian notion of ‘using someone merely as a means.’ At first sight, this may seem to be an unpromising strategy. In her classic article ‘The Trolley Problem,’ Judith Jarvis Thomson argues, on the basis of a scenario known as the ‘Loop’ case, that a solution in terms of this Kantian notion encounters insoluble difficulties (Thomson 1985: 1401–3). Her argument is widely considered convincing. Against this view, I argue that there is a plausible conception of ‘using merely as a means’ that yields an explanation of the moral difference between the ‘Bystander’ and the ‘Footbridge’ case without foundering on the Loop case.

I do not provide a *justification* of the moral prohibition on using persons merely as means, since this requires separate discussion, nor do I fully settle the issue of whether diverting the trolley in the Bystander and the Loop case is morally permissible. Thus, I do not claim to deliver a full solution to the Trolley Problem if this is thought to require a ‘satisfying justification for our initial intuitive judgments’ (Hurka 2016: 150).<sup>4</sup> Instead, I use a different standard for what counts as a solution, also common in the literature, according to which a solution requires a principled account of the *salient moral difference* between the Bystander and the Footbridge case. Thomson claims that the Kantian prohibition on ‘using someone merely as a means’ cannot provide such an account; my aim in this essay is limited to showing how that challenge can be met.

<sup>2</sup> Some authors explain the so-called ‘deontological’ responses to trolley dilemmas as resulting from emotional reactions (Greene et al. 2001). Others argue that our intuitive moral heuristics become unreliable when applied to artificial hypothetical situations such as trolley scenarios (e.g., Gigerenzer 2008: 11; Appiah 2008: 96–101; Bruers 2016).

<sup>3</sup> Several authors criticize trolley dilemmas for abstracting from the details and uncertainties that characterize real-life situations of moral decision making. For different versions of this objection, see, e.g., Wood (2011), Fried (2012).

<sup>4</sup> I do not mean to rule out that readers who endorse the methodology of reflective equilibrium may come to believe that I accomplish all of this if they find the argument of this chapter convincing. But my aim is more limited.

I begin by presenting the Trolley Problem, explaining Thomson's challenge to a solution in terms of the notion of 'using' someone 'merely as a means' and specifying the desiderata for an account that can meet this challenge (Section 1). I then offer an account of 'using' someone 'merely as a means' (Section 2). I argue that this account enables a solution to the Trolley Problem by explaining the morally salient difference between the Bystander and the Footbridge case (Section 3), and I show that it can also handle the Loop case (Sections 4 and 5).

## 1. The Trolley Problem and the Loop Challenge

As is by now familiar, Thomson articulated the Trolley Problem in a discussion of the following scenario devised by Philippa Foot:

[Someone] is the driver of a runaway tram which he can only steer from one narrow track on to another; five men are working on one track and one man on the other; anyone on the track he enters is bound to be killed. (Foot 1978 [1967]: 23)

Foot examined the moral difference between a driver's steering towards the one man in this case and the case of a mayor who frames an innocent person to prevent riots, or a transplant surgeon who kills a random healthy person in order to use his organs to save a greater number of patients. Problematising some of Foot's core claims, and in order to focus on the *structure* of the moral issue without introducing complications arising from extraneous differences between the scenarios (such as differences having to do with what is specific to organ transplantation as compared to riot prevention or trolley malfunction), Thomson created an entire family of trolley cases (Thomson 1985; cf. Thomson 1976).

In one variant—the Bystander scenario—the original situation is the same as in the Driver case described by Foot, except that the driver is incapacitated and a bystander is standing near a switch that will allow him to divert the trolley onto the track with one workman on it. In another variant—the Footbridge case—someone is standing on a footbridge over the track and sees the trolley hurtling towards five workmen. Standing

nearby is a 'really fat man' who is leaning over the railing.<sup>5</sup> Pushing him would cause him to fall into the path of the oncoming trolley, which would kill him but stop the trolley and save the five (and the agent cannot stop it by jumping in front of it; Thomson 1985: 1409). There has been a proliferation of sometimes unusual scenarios in the discussion to date, such as cases involving people stuck on oversized Lazy Susans (Kamm 2007) and heavy men on remote-controlled roller skates (Unger 1996: 90). But there is nothing outlandish about the basic philosophical question behind the original scenarios on the basis of which Thomson formulated the Trolley Problem:<sup>6</sup>

In both cases, one will die if the agent acts, but five will live who would otherwise die—a net saving of four lives. What difference in the other facts of these cases explains the moral difference between them?

(Thomson 1985: 1396, cf. 1401)

One of the possible answers Thomson considers but rejects is the reply that the agent on the footbridge uses the heavy man merely as a means, whereas the agent in the Bystander scenario does not. She allows that on the face of it, the Kantian notion of 'using' someone 'merely as a means' may seem to capture the moral difference between the cases. As she puts it, the bystander does not need or use the one workman to save the five, because the latter's presence on the track contributes nothing to achieving the agent's end of saving the five. By contrast, the agent on the footbridge does need to use the one if he is to save the five (Thomson 1985: 1402).

Thomson argues that the Kantian notion of 'using merely as a means' does not provide a satisfactory answer, however, and to this end she introduces a modified Bystander scenario called the Loop case. As in the Bystander case, the tracks diverge, but now they reconnect further along to form a loop. The one workman, who is stuck on the track beyond

<sup>5</sup> Given the negative stereotypes often associated with being 'fat,' Thomson's description may prejudice responses to the scenario. Hence, the man on the bridge is nowadays sometimes cast as carrying a backpack of trolley-stopping proportions, or simply as 'heavy.'

<sup>6</sup> In a more recent essay, Thomson argues that the Trolley Problem is not really a problem after all, because diverting the trolley and pushing the man are both impermissible (Thomson 2008). I do not find her new argument convincing, for reasons including but not limited to the problems mentioned by FitzPatrick (2009).

the fork, is so heavy that his body would stop the trolley, whereas the five are so thin that all five (and only the five) would be killed if the trolley were to hit them. The trolley is heading towards the five, but if the bystander diverts it towards the one workman, this will save the five and kill the one instead. In this Loop scenario, Thomson claims, the bystander needs the one workman to save the five in the same way that the agent on the footbridge needs the heavy man to stop the trolley. This suggests, she claims, that diverting the trolley towards the one to save the five in the Loop scenario is a case of ‘using him merely as a means’ (Thomson 1985: 1403).

Morally speaking, this would put the Loop case on a par with the Footbridge case: in both scenarios, the heavy man would be used merely as a means. As far as the details of the scenario are concerned, however, the Loop scenario strongly resembles the Bystander case. The only difference is a connecting bit of track in the distance. Thomson argues that it is implausible to assume that this connecting bit of track creates a major moral difference between the Bystander and the Loop scenario. Moreover, she asserts, given the similarities with the Bystander case, it still seems permissible for the bystander to divert the trolley in the Loop case, even though this would involve using the one workman merely as a means. If it is *permissible* in the Loop case to divert the trolley despite the fact that this involves using the heavy man merely as a means, however, then what makes it *impermissible* in the Footbridge case to push the heavy man in front of the trolley cannot be the fact that this involves using him merely as a means. This in turn implies that the prohibition against using someone merely as a means cannot explain the moral difference between the Bystander and the Footbridge case.

Thomson therefore concludes that the Kantian notion is incapable of solving the Trolley Problem:

[T]here is no plausible account of what is involved in, or what is necessary for, the application of the notions ‘treating a person as a means only’, or ‘using one to save five’, under which the surgeon [or the agent who pushes the heavy man off the footbridge] would be doing this whereas the agent in this variation [i.e., the bystander who diverts the

trolley in the Loop scenario]...would not be. If that is right, then appeals to these notions cannot do the work being required of them here. (Thomson 1985: 1403)

Some Kantians have replied that Thomson's intuition is mistaken. Samuel Kerstein has claimed as much, arguing that diverting the trolley in the Loop case is impermissible because it involves using the one merely as a means (Kerstein 2013: 124). This line of response, however, is unlikely to satisfy those who, like Thomson herself in her 1985 paper and again in her most recent discussion of the issue (2016: 128–32), are wondering just *how* the addition of a bit of track causes such a massive moral difference. We can press the issue by assuming that the additional bit of track includes a drawbridge in the looping section between the one and the five. If one assumes that diverting the trolley would involve using the heavy man merely as a means in the Loop case but not in the Bystander case, one must also maintain that, in this modified Loop scenario, diverting the trolley would involve using the workman merely as a means when the distant drawbridge is down but not when it is up (assuming the other details of the case stay the same). This result may seem far-fetched, and Thomson indeed says that it would be 'thoroughly odd' if diverting the trolley turned out to be right or wrong depending on a loop connection that may or may not obtain (Thomson 2016: 131). Moreover, given that the bridge is located *beyond* the heavy workman and his body will stop the trolley no matter whether the bridge is open or closed, it is unclear *how* the position of the bridge could affect the permissibility of diverting the trolley. A reply to Thomson's challenge should address the Loop scenario's similarity to the Bystander case and explain the significance of the extra bit of track.

A full reply to Thomson's challenge would therefore need to include at least the following elements:

1. a plausible conception of 'using someone merely as a means';
2. an indication of the salient moral difference between the Bystander and the Footbridge case in terms of this notion—that is, a basic solution to the Trolley Problem; and

3. a successful extension of this basic solution to the Loop case, including an explanation of the significance (or lack thereof) of the extra bit of track.

In the remainder of this chapter, I develop these elements in turn, starting with the notion of ‘using someone merely as a means.’

## 2. ‘Using Someone Merely as a Means’

In the Kantian literature, the prohibition on ‘using’ someone ‘merely as a means’ is commonly explained in terms of the violation of a requirement concerning the *consent* of the person who is used as a means. This requirement makes it possible to distinguish between abuse and permissible use. There are various accounts of the type and object of the required consent,<sup>7</sup> but it is not necessary to discuss alternative accounts here. My aim is limited to showing that there is at least one plausible understanding of the Kantian notion of ‘using someone merely as a means’ that can meet Thomson’s challenge. The conception of ‘using someone merely as a means’ that I use for the purpose of this chapter is the following:

USING MERELY AS A MEANS: an agent uses another person ‘merely as a means’ if and only if (1) the agent *uses* another person *as a means* in the service of realizing her ends (2) without, as a matter of moral principle, making this use *conditional* on the person’s *consent*, where (3) the required consent is *actual* genuine consent to being used by the agent in a particular manner, as a means to the agent’s end.

I have argued elsewhere (Kleingeld 2020), with the necessary exegetical details, that this account of the prohibition is plausible as an interpretation of Kant’s texts, and indeed more plausible than alternative interpretations. In the present essay, I bracket exegetical debates. Because the canonical exposition of the prohibition on using others ‘merely as means’ is found

<sup>7</sup> See, e.g., Kerstein (2013), Korsgaard (1996), O’Neill (1989), Parfit (2011), Wood (2008).

in Kant's work, however, I will mention a few key passages as I present the three elements of the account.

## 2.1 The Use-as-Means Condition

The classic formulation of the prohibition on using persons 'merely as means' is found in Kant's Formula of Humanity in his *Groundwork for the Metaphysics of Morals*:

So act that you use the humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means. (G 4: 429)<sup>8</sup>

Note, first, that Kant speaks of 'using' rather than of 'treating' someone merely as a means. In current debates, it is customary also to speak of 'treating' a person merely as a means, and even to use these terms alternately. 'Treating' has a much broader scope than 'using,' however. It also includes mere attitudes toward others that do not amount to any actual *use* of them as means, such as regarding others as being of potential use.<sup>9</sup> 'Treating' seems to have been introduced into the discussion by influential but infelicitous translations of the Formula of Humanity.<sup>10</sup> Kant's terminology is that of 'using' persons (merely) as means. On the account I shall be employing in this chapter, the first condition without which we cannot properly say that an agent uses a person 'merely as a means' is that the agent indeed *uses* the person *as a means* to her end.

<sup>8</sup> Kant's writings are cited by the abbreviated title, using the Akademie volume and page numbers (Kant 1900–). Translations are found in the Cambridge Edition of the Works of Kant (Kant 1992–2016). Abbreviations: CLMP = Collins Lectures on Moral Philosophy; CPR = *Critique of Practical Reason*; FLNL = Feyerabend Lectures on Natural Law; G = *Groundwork for the Metaphysics of Morals*; MM = *Metaphysics of Morals*; MM Vig = Lectures on the Metaphysics of Morals Vigilantius.

<sup>9</sup> Treating others merely as means does violate the demands of the Formula of Humanity, but it involves a different type of violation: the agent fails to promote their humanity at the same time as an end, as Kant puts it (G 4: 430).

<sup>10</sup> Kant uses '(ge)brauchen,' which means 'use,' but both H.J. Paton and Lewis White Beck use 'treat' in their translations. Kant uses 'treating' at least once ('behandeln,' G 4: 433). This is compatible with 'using' being the dominant term, since 'using' is a form of 'treating.' But since not all 'treating' is 'using,' I employ Kant's terminology of 'using' when discussing the Formula of Humanity.



There is another terminological issue that requires comment at the outset, namely the appropriateness of speaking of using ‘a person,’ given that the Formula of Humanity speaks of using ‘the humanity in a person’ merely as means. By ‘humanity’ Kant understands that which distinguishes humans from other animals, and this is their capacity to act in accordance with rational principles and to set ends independently of inclination (G 4: 428–31, 434). He argues, however, that in a human person this capacity is inseparably connected with the body and that using part of an entity constitutes using that entity (MM 6: 279, CLMP 27: 387). In other words, using a person’s body or using a person’s rational capacities counts as using a person. Similarly, when using a particular program on my computer, I can say that I am using my computer. Kant indeed often speaks of ‘using a *person* merely as a means’ or ‘using *another human being* merely as a means’ (G 4: 429), and in this chapter I will do the same.

But when does an agent use another person *as a means*? This question can be answered in two different ways. We can describe the agent’s action either from the agent’s own perspective, in terms of her end, the means by which she intends to promote or achieve it, and her associated line of reasoning; or we can describe the action from an external narrator’s perspective, focusing on the agent’s observable behavior and another person’s causal contribution to the realization of the agent’s presumed end.

Authors discussing the Trolley Problem often describe the agent’s action from an external perspective. In order to determine whether the one person is used merely as a means, they examine whether that person’s presence is required for (or contributes to) achieving the agent’s end. Thomson, for example, suggests that the one workman is used merely as a means ‘only if, had the one gone out of existence just before the agent started, the agent would have been unable to save the five’ (Thomson 1985: 1402). Since this criterion is external to the agent’s own means–ends reasoning, it may describe an agent as using another person merely as a means even if that person does not serve *as a means* in the agent’s own practical reasoning at all.

Kant, by contrast, takes the agent’s own practical reasoning to be the relevant point of reference for determining whether the agent is using

another person as a means. In the *Groundwork*, his examples of agents who use others merely as means are phrased in terms of the agent's practical reasoning: in terms of what the agent 'has in mind' or 'wants' and what 'principles' he is acting on (G 4: 429–30).

For the purposes of this chapter, I will similarly take the relevant point of reference to be the agent's own practical reasoning. An agent uses another person as a means if and only if she wants to reach a certain end, believes that she can reach or further this end mediately, by using another person as a means, and uses the person for the sake of reaching or furthering her end.

## 2.2 The Consent Proviso

Not every case in which an agent uses someone *as a means* also involves the agent's using someone '*merely*' *as a means*. We need a second condition to pick out the proper subset, and this is the role of the consent requirement. On the account I will be employing in this chapter, an agent who uses someone as a means uses this person '*merely*' as a means if and only if she fails to honor the moral constraint that her use of others must be *conditional* on their *consent* (and the relevant type of consent is actual consent, as explained in Section 2.3). Making one's use of others conditional on their consent involves committing oneself not to use others as means unless they consent to being used as such. The agent ought to honor this constraint as a matter of moral principle, that is, as a matter of recognizing others as 'ends in themselves' (as Kant puts it), rather than as a matter of following mere personal inclination.

Kant formulates the consent proviso in terms of a limiting condition on the agent's maxims. Maxims, in his moral theory, are the major premises in the agent's practical reasoning: the principles on the basis of which the agent acts (G 4: 421n.; CPrR 5: 19). Although the Formula of Humanity does not contain an explicit reference to the agent's maxims, Kant repeatedly emphasizes that it articulates a limiting condition on maxims. He claims that 'the formula says' that rational beings 'must serve *in every maxim* as the *limiting condition* of all merely relative and

elective ends' (G 4: 436, emphasis added; also 437). He rephrases this idea by stating that rational beings must serve as the 'highest limiting condition in the use of all means' (G 4: 438).

Kant explains the nature of this limiting condition in terms of the *consent* of the person who is used as a means. In the *Groundwork*, he gives the example of a person who wants to obtain money by getting a loan with a false promise to repay it. He explains that the false promisor wants to use the lender 'merely as a means' by saying that the lender 'cannot possibly *consent* to [the false promisor's] way of behaving toward him' (G 4:429). This suggests that if an agent uses another as a means to her end without having made this use conditional on the other's consent, then she uses the other *merely* as a means.

### 2.3 The Relevant Type of Consent Is Actual Consent

In the literature, the relevant type of consent is variously conceived as hypothetical, rational, possible, or actual consent. The textual evidence suggests, however, that the relevant type of consent is actual consent. Kant argues that actual consent is required whenever one person wants to use the property or powers of another (MM 6: 285, 271–2). As an example, he mentions that the use of servants requires their *actual* consent: a servant 'must also be an end...and not merely a means. He must also want it' (FLNL 27: 1319; cf. MM 6: 283). Furthermore, Kant introduces the loan example in the *Groundwork* by saying that the lender does not actually consent to the use the false promisor wants to make of him ('without [the lender] at the same time containing in himself the [false promisor's] end,' given that the latter misrepresents his real end, G 4: 429).

For a full explanation of the requirement that agents use others as means only on condition that the others actually consent, much more would need to be said, of course, but for the purposes of this chapter it is not necessary to do so. None of the agents in the trolley cases under consideration obtain the actual consent of the men who are about to be run over by the trolley. If these agents use the men in these scenarios as means, they clearly do so regardless of whether the men actually consent, so a discussion of the finer details of applying this requirement is unnecessary.

Is this account a *plausible* conception of ‘using someone merely as a means,’ as Thomson requires? It straightforwardly explains the notion in terms of its constituent elements, and it distinguishes effectively between using a person ‘*merely as a means*’ and using a person ‘as a means, but *not merely as a means*.’ Furthermore, it matches widely shared moral convictions concerning the requirement of informed and voluntary consent in cases of using others as means. For example, informed consent requirements in medical research contexts impose on researchers an obligation to make their use of human test subjects conditional on those subjects’ actual consent (to being used in the manner specified and in the service of a particular research project). This widely shared moral conviction matches the Kantian prohibition as I have described it here. Thus, this conception of the prohibition on using others ‘merely as a means’ seems to have enough initial plausibility, even without a full defense of the prohibition as such (which, as mentioned, it is not the aim of this chapter to provide), to warrant examining its potential for handling the Trolley Problem.

### 3. The Basic Solution to the Trolley Problem

Returning now to the trolley scenarios, we can draw out important implications of this account of ‘using merely as a means’ for the assessment of the cases at issue. Clearly, on this conception of the notion, whether the agent is using the workman ‘merely as a means’ cannot be established by examining the causal relation between the death of the one man and the survival of the five. We need to focus our attention *on the agent*, namely on whether she *uses* the one *as a means* without making this use conditional on the workman’s *actual consent*.

None of the agents in the Bystander, Footbridge, and Loop cases have obtained the actual consent of any of the men involved. This entails that if the agent uses the one man *as a means* to her end of saving the five, then she uses him *merely as a means*. The crucial question in each of the three trolley scenarios, therefore, is whether the agent *uses* the one man *as a means*, and this, as noted above, depends on her practical reasoning.

In what follows, I first show how the proposed account of ‘using merely as a means’ explains the morally significant difference between

switching the trolley towards the track with the one workman (in the Bystander case) and pushing the heavy man off the bridge (in the Footbridge case). I then turn to the Loop case in the next section.

As initially suggested by Thomson, and as widely agreed in the Trolley Problem literature, the agent who diverts the trolley in the Bystander scenario does *not* use the one workman *as a means* (e.g., Thomson 1985: 1402).<sup>11</sup> The agent is assumed to have the end of saving more rather than fewer lives. She finds herself forced to choose: either she lets the trolley continue, in which case five workmen die, or she diverts it, in which case one workman dies. She realizes that she will achieve her end *by steering* the trolley onto the sidetrack, and this is why she decides to divert it. The fact that the workman dies as a consequence of her diverting the trolley does not mean that she uses him as a means to her end. At no point does he enter into her practical reasoning as a means to her end of saving more rather than fewer lives. In fact, she could save even more lives if he were not there. Since the agent who diverts the trolley in the Bystander scenario does not use the one workman *as a means*, this trivially entails that she does not use him *merely* as a means.

In the Footbridge case, by contrast, an agent who deliberately pushes the heavy man off the bridge as a means to her end of saving the five does use him *merely* as a means. Here the conditions for using a person ‘merely as a means’ are satisfied. First, it is clear that the agent *uses* the heavy man *as a means*. Her reason for pushing him is that his body will stop the trolley and she will thereby reach her end of saving the five. In fact, she regards the heavy man as ‘the one’ (as compared to ‘the five’ on the tracks) *because* she can use him to stop the trolley. Given his weight, his expected impact on the trolley, and her own ability to push him over the edge, she regards the heavy man as ‘the one’ qua potential trolley stopper. If he were thin, or if he could not be moved, she would not consider pushing him. Furthermore, if there were a thin man standing next to the heavy man on the bridge, *only* the heavy man could be used as a means to stop the trolley, and he would suffice. Any agent who

<sup>11</sup> As is customary in the debate, I assume in the discussions to follow that the agent is neither confused, nor acting on an independent desire to kill one workman, nor acting on mistaken beliefs (including the mistaken belief that the five are saved by killing the one workman, rather than by diverting the trolley).

wants to save more lives rather than fewer would therefore face only the question of whether to push *him*—not whether to push ‘the two.’ Thus, the heavy man’s role as ‘the one’ in the Footbridge scenario is defined by his potential use as a means to stop the trolley. And if the agent pushes him in order to stop the trolley so as to save the five, she necessarily uses him as a means.<sup>12</sup>

Second, the details of the Footbridge scenario indicate that the agent who uses the heavy man as a means to stopping the trolley does so without making her use of him conditional on his actual consent. Because she deliberately uses him as a means to stopping the trolley without having made this conditional on his actual consent, she uses him *merely as a means*.

The basic solution to the classic Trolley Problem in terms of this account of ‘using merely as a means,’ then, is as follows. The agent in the Bystander scenario who diverts the trolley to save five does not use the one workman as a means to her end of saving the larger number. Since the use-as-means condition is not satisfied, she does not use the one workman merely as a means. By contrast, the agent in the Footbridge scenario who deliberately pushes the heavy man necessarily uses him as a means, and given the details of the scenario, she does so without making her use of him conditional on his consent. Thus, she necessarily uses him merely as a means.

In this way, the proposed conception of ‘using merely as a means’ yields an account of the salient moral difference between the Bystander and the Footbridge case: the agent in the first case does not use the man merely as a means, whereas the agent in the second case does.

On the assumption that it is morally impermissible to use another merely as a means, this establishes that pushing the man off the footbridge is impermissible, but it does not establish that it is morally permissible—let alone required—for the bystander to divert the trolley. This is because the argument above does not rule out that there are other considerations that render it impermissible. When a certain action can

<sup>12</sup> This analysis of the Footbridge scenario does not change if we recast the heavy man as a thin man strapped to a backpack (or to a wheelchair, a voluminous suit, or any other item) of trolley-stopping proportions. If it is impossible to toss only his backpack off the bridge, then pushing the backpack to which he is attached counts as pushing him (see Section 2.1 above).

be performed *only* by using someone merely as a means, as in the Footbridge case, we can indeed conclude that the principle implies that the action is impermissible. But it does not follow that any action that is performed *without* using a person merely as a means is therefore permissible. The Formula of Humanity not only prohibits using persons merely as means but also requires promoting them as ends in themselves. Agents who avoid using anyone merely as a means may still be violating this further requirement (see Kant's *Groundwork* example of the person who refuses to help others in need). For the purpose of this chapter, however, it is not necessary to pursue such cases.

#### 4. Two Construals of the Loop Case

The above account of 'using someone merely as a means' has an implication that is crucial for the analysis of the Loop case: one and the same action does or does not constitute an agent's using someone 'merely as a means' depending on the agent's underlying practical reasoning. This implication does not rise to the surface in the discussion of the Bystander and Footbridge cases. Given the details of these scenarios and the reasoning attributed to the agents, the agent in the Footbridge case *does* use the one man merely as a means, but the agent in the Bystander case *does not*. I will argue that in the Loop case, by contrast, we can ascribe different lines of practical reasoning to the agent, such that the agent *does or does not* use the heavy workman merely as a means, depending on what we stipulate. This will make it possible to formulate the response to Thomson's challenge in the next section.

It is not difficult to conceive of a bystander in the Loop case who uses the heavy man merely as a means. Suppose an agent named Tili is standing near the switch. Tili sees the trolley approaching the five, notices the heavy man on the track beyond the fork, and recognizes that she could use him to prevent the trolley from killing the five. She realizes that her situation is such that she has to make a decision without having any

further knowledge about the six people involved. I abbreviate the description of this aspect of the situation as ‘S.’<sup>13</sup> She reasons as follows:<sup>14</sup>

TILI: When forced to choose in S, I will save more rather than fewer human lives, even if this involves my using others as means to this end without their actual consent. Now I am forced to choose in S: either I let the trolley continue towards the five, which will save one life, or I divert it towards the heavy man so as to use him as a means to stopping the trolley, which will save five lives. Five is more than one, so I shall save the five and divert the trolley in order to use the heavy man as a means to stopping it.

On the basis of this line of reasoning, Tili diverts the trolley towards the heavy man. In situations like these, she reasons from the general action principle stated in the first sentence (‘when forced...’), and if she ever found herself in the Footbridge scenario, this principle would lead her to push the heavy man. Reasoning now from this principle in the Loop case, she decides to divert the trolley towards him in order for his body to halt it. This means that she uses the heavy man as a means to her end, and she also does so without his actual consent. Thus, she uses the heavy man *merely as a means* in the full sense of the account presented above.

The details of the Loop scenario, in particular the scenario’s structural similarities to the Bystander case, also make it possible, however, for us to imagine a bystander who diverts the trolley *without* using the heavy man *as a means*, and hence without using him *merely as a means*.

Suppose an agent named Manuela<sup>15</sup> reasons from an action principle that includes the morally required limiting condition: ‘When forced to

<sup>13</sup> This is meant to bracket those complications that the trolley scenarios were specifically designed to bracket, such as special relationships between the agent and the workmen, specific characteristics of some of the workmen involved, etc.

<sup>14</sup> The focus here is on patterns in moral reasoning and justification, not psychology. This is why I abstract from the psychological fact that most real agents would feel conflicted, panicked, and horrified.

<sup>15</sup> I introduce a second agent (rather than two possible patterns of reasoning for one agent) so as to avoid any suggestion that the agent is ‘tinkering’ with her maxims to produce a morally



choose in S, I will save more rather than fewer human lives, provided I do not use anyone as a means to this end without their actual consent.' Unlike Tili, Manuela would refrain from pushing the heavy man in the Footbridge scenario. But what about the Loop case?

Manuela knows that if she lets the trolley continue, five people will die. She also knows that if she diverts the trolley, it will hit and kill the heavy man and stop there. That is, she knows that if she diverts the trolley, this will lead to one death. She reasons as follows:

MANUELA: When forced to choose in S, I will save more rather than fewer human lives, provided I do not use anyone as a means to this end without their actual consent. Now I am forced to choose in S: either I let the trolley continue towards the five, which will save one life, or I divert it towards the heavy man, which will kill him but save five lives. Five is more than one, so I shall save the five by diverting the trolley.

On the basis of this line of reasoning, Manuela diverts the trolley. Her end is to save more rather than fewer human lives. She knows that if she diverts the trolley, one person dies, and that if she lets it continue, five die. She will achieve her end by diverting the trolley to the side *where only one man will die*, and this is why she diverts it. Her reasoning includes the morally required proviso, and at no point does the heavy man enter into her reasoning *as a means*. It is not that she neglects to take note of important facts about the case: she knows that his body will stop the trolley and that this will kill him. But this prediction as such does not include any instrumental role for him, and she could have made it even if she had not yet seen the five ahead. Since she does not use the heavy man as a means, she does not use him *merely* as a means. So in terms of the above account, Manuela diverts the trolley *without* using the heavy workman merely as a means.

There is a related context in which similar distinctions are drawn. This is the discussion of whether a bystander in the Loop case who

desirable result. I stipulate that Tili and Manuela really have adopted their respective action principles and really are committed to acting accordingly in the relevant circumstances.

decides to let the trolley *continue* towards the five would be *using the five as means* to save the one—say, in acting on the principle that it is better to do nothing than to kill one. Commentators generally deny that this inactive bystander would be using the five as means, for the very same reason that I have offered in support of the claim that Manuela does not use the heavy man as a means. They emphasize that the five do not enter into the inactive bystander's practical reasoning *as means*.<sup>16</sup> They say that the fact that the five die as a *consequence* of the bystander's decision to let the trolley continue does not entail that she uses them *as means*. By parity of reasoning, then, the fact that the heavy workman dies as a *consequence* of Manuela's diverting the trolley does not entail that she uses him *as a means*, let alone that she uses him *merely* as a means. Her end is to save more rather than fewer human lives, and she attains this end *by diverting* the trolley to the side where—as she predicts—one man will die.

The similarities between the Loop and the Bystander case make it possible to conceive of an agent who diverts the trolley in the Loop case *without* making instrumental use of the heavy man. An important difference between the Bystander and the Loop case is, of course, that in the Loop scenario the agent would not be able to save the five if the one workman were not there. But the truth of this counterfactual does not invalidate the analysis of the actual case at hand.

The significance of the extra bit of track, then, is that it makes it possible to construe the agent's reasoning in the Loop scenario as similar to that in either the Bystander case or the Footbridge case. This makes the Loop case fundamentally different from both. In the Bystander scenario, the agent who diverts the trolley does not use the one workman as a means, and hence she does not use him merely as a means. In the Footbridge case, the agent who deliberately pushes the heavy man uses him merely as a means. In the Loop scenario, by contrast, the agent who diverts the trolley towards the heavy man does (e.g., Tili) or does

<sup>16</sup> For example, Kerstein (2013: 123) writes: 'that you allow the five to be killed does not entail that you are using them.'

not (e.g., Manuela) use him merely as a means, depending on their line of reasoning.<sup>17</sup>

The fact that the Loop case allows for such different lines of reasoning explains why the scenario has generated such radically different moral assessments in the literature. Some authors, including Thomson in her 1985 essay, assimilate the Loop case to the Bystander scenario and regard it as permissible to divert the trolley towards the one. Others, including Kerstein (2013) and Michael Otsuka (2008), assimilate the Loop case to the Footbridge scenario and regard it as impermissible to divert the trolley. The argument of this section suggests that it is a mistake to assume that the Loop scenario should be assimilated to only one of the two. The scenario permits both construals. Thus, the account of ‘using merely as a means’ applied in this chapter helps to explain why the Loop scenario generates so much disagreement, and I take this to be one of its strengths.

The main result of this section, however, is the conclusion that it is possible for a bystander to divert the trolley in the Loop case without using the heavy man as a means, and hence without using him merely as a means. According to Thomson, quoted above, on any plausible account of what is involved in ‘using someone merely as a means,’ any bystander who diverts the trolley in the Loop scenario would be doing so. The argument of this section shows that this is not true, and this makes it possible to respond to Thomson’s Loop challenge.

## 5. Extending the Basic Solution to the Loop Scenario

Recall that Thomson’s challenge is as follows: if it is permissible to divert the trolley in the Loop case, even while using the heavy man merely as a means, then what makes it impermissible to push the heavy man in the

<sup>17</sup> One might wonder whether it is possible to generate different construals for the Bystander case, too. I do not think it is, at least if we assume that the Bystander does not act on the basis of mistaken beliefs (see note 11). But even if it were possible, then this would not be a problem for the argument of this chapter. The Trolley Problem would then require a solution in terms of the salient moral difference between the Footbridge case and the two Bystander cases (in its classic and loop versions). That is, the solution would then consist in the salient moral difference between the Footbridge and the Loop case described in Section 5.

Footbridge case cannot be the fact that this involves using him merely as a means. This in turn implies that the Kantian prohibition on using someone ‘merely as a means’ cannot solve the Trolley Problem. In the previous section I argued that diverting the trolley in the Loop scenario does *not necessarily* involve using the heavy man merely as a means. But what difference does that make? If Tili diverts the trolley, she uses the heavy man merely as a means, and if it is nevertheless permissible *for her* to divert the trolley—as Thomson (1985) assumes it is—then the Loop case still seems to undermine the basic solution of the Trolley Problem in Section 3 above.

In order to answer Thomson’s Loop challenge, we need to look more closely at the relation between the permissibility of an action and the practical reasoning that underlies it. I will argue that the Loop challenge can be answered by combining the following two claims, both of which are familiar Kantian principles: (1) A morally permissible action remains permissible even when it is performed on the basis of a morally impermissible action principle; and (2) if an action can be performed only on the basis of morally impermissible action principles, then the action is impermissible.

In the previous section I showed that Tili and Manuela can divert the trolley on the basis of practical reasoning of very different moral quality. It is a familiar idea within Kantian ethics that a permissible action can be performed on the basis of permissible or impermissible action principles (‘maxims’). A helpful illustration of this point is Kant’s well-known *Groundwork* example of a shopkeeper who charges children the right price on the basis of an impermissible action principle, namely on the basis of the maxim of only pursuing his own long-term interests (G 4: 397).<sup>18</sup> Charging children the right price (the action), Kant says here, is ‘in accord with duty,’ regardless of whether the shopkeeper’s maxim satisfies moral requirements. The action is required by the duty of honesty, and a shopkeeper who charges children the right price still performs an action that is ‘in accord with duty,’ even if his practical reasoning is guided by the principle of self-interest. To add another example: resuscitating the victim of an accident does not become impermissible if the

<sup>18</sup> Kant argues elsewhere that acting on the maxim of self-interest is impermissible (MM 6: 451; cf. also MM Vig 27: 621).

agent uses the victim merely as a means to practicing her life-saving skills, on the basis of an egoistic principle (say, in order to get a higher salary after she passes a life-saving test, and she would not resuscitate anyone if it weren't instrumental to reaching her goal of increasing her salary). If the duty to help others in need requires that she resuscitate the victim, then she does the right thing, even if she does it for the wrong reason.

Thus, a morally permissible or required *action* does not become impermissible when performed by an agent who reasons from an impermissible *maxim*. Nor does an impermissible maxim become permissible if it happens to lead an agent to perform a permissible action. The distinction between the moral status of the action and the moral status of the underlying maxim (and the associated reasoning) is crucial to Kantian moral theory.<sup>19</sup>

This distinction explains how the permissibility of diverting the trolley in the Loop scenario is compatible with the impermissibility of using the heavy man merely as a means. The permissibility of diverting the trolley concerns the *action* (in abstraction from the reasoning that underlies it), whereas the impermissibility of using the heavy man merely as a means concerns the agent's underlying *reasoning*. Tili and Manuela perform *the same action* in the Loop case: they both divert the trolley. The permissibility of diverting the trolley does not vary with the different moral quality of their practical reasoning. If diverting the trolley is permissible, it is permissible for Tili too, even though she uses the heavy man merely as a means in the process. And this clearly does *not* mean that it is permissible for her to use the heavy man merely as a means.

We are now in a position to answer the question that still remains: suppose that diverting the trolley in the Loop case is permissible, and hence permissible for Tili, then how can the prohibition against using others merely as means explain the impermissibility of pushing the

<sup>19</sup> The terminology with which authors draw related distinctions varies. For example, Christine Korsgaard distinguishes between 'act' and 'action' (Korsgaard 2009: 12), and T.M. Scanlon distinguishes between the 'permissibility' and the 'meaning' of an action (Scanlon 2008: 89–121, esp. 101). Non-Kantian authors have drawn similar distinctions (e.g., Philip Pettit 2015, ch. 5).

heavy man off the footbridge? The answer turns on a crucial difference between the Footbridge and the Loop case.

In the Footbridge case, the action of pushing the heavy man off the bridge (given the details of the case) *necessarily* involves the agent's using him merely as a means, as I argued in Section 3. For an action to be morally permissible, it should of course at least be possible for an agent to perform it without violating moral constraints. Thus, if an action can be performed only on the basis of morally impermissible action principles, then it is impermissible. This is precisely the situation in the Footbridge case, since pushing the heavy man necessarily violates the prohibition on using others merely as means (given the details of the scenario and the kind of reasoning that can be attributed to the agent given these circumstances). Pushing the heavy man in front of the trolley is therefore impermissible, and what explains its impermissibility is indeed the fact that it (necessarily) involves using him merely as a means.

In the Loop case, by contrast, diverting the trolley does *not necessarily* involve the agent's using the heavy man merely as a means, as I argued in Section 4. Thereby the Loop scenario opens up the possibility that diverting the trolley is permissible or even morally required. And if diverting it is indeed permissible (which is an issue I do not settle in this essay), then diverting is permissible even for an agent such as Tili who uses the heavy man merely as a means. She would perform a permissible *action*, even though she would do so on the basis of an impermissible *action principle* (*maxim*).

In sum, despite initial appearances to the contrary, the following two claims are compatible: (1) diverting the trolley in the Loop case is permissible, even for an agent who uses the heavy man merely as a means in doing so; (2) what makes it impermissible to push the heavy man in the Footbridge case is the fact that this (necessarily) involves using him merely as a means. The Loop case does not undermine the basic solution to the Trolley Problem after all.

It seemed to Thomson, in her 1985 paper, that diverting the trolley in the Bystander and the Loop case is permissible. It is not the aim of this chapter to determine fully whether diverting in either case really is permissible, but it seems to me that Thomson's impression was correct

and, moreover, that Manuela's reasoning stands a good chance of satisfying Kantian criteria. For reasons mentioned above, however, a full determination of what is permissible in these cases requires more argument, to assess whether the agent would violate moral requirements other than the one discussed here.<sup>20</sup> It is not necessary to settle this for the more limited purpose of this essay. My aim here has been to offer an explanation of the *salient moral difference* between the Bystander and Footbridge case that does not founder on the Loop case. I have argued in defense of the following account of the relevant difference: the agent in the Footbridge case who pushes the heavy man off the bridge to stop the trolley necessarily uses him 'merely as a means,' whereas the agent who diverts the trolley in the Bystander case does not use him 'merely as a means.' Regardless of whether diverting the trolley in the Bystander case turns out to be permissible, all things considered, this difference will remain.<sup>21</sup>

## References

- Appiah, Kwame Anthony. 2008. *Experiments in Ethics*. Cambridge, MA: Harvard University Press.
- Bruers, Stijn. 2016. "Can Deontological Principles Be Unified? Reflections on the Mere Means Principle." *Philosophia* 44: 407–22.
- FitzPatrick, William J. 2009. "Thomson's Turnabout on the Trolley." *Analysis* 69: 636–46.

<sup>20</sup> Further issues that deserve consideration at that point include the distinction between killing and letting die, and the question of whether killing one is morally worse than 'doing nothing' even if doing nothing means letting five people die. I doubt that it follows from Kantian principles that killing one is morally worse than letting five die, assuming that no person is used merely as a means in the process—especially if we take pains to avoid loading the term 'killing' with moral content in a question-begging way. But pursuing this issue lies beyond the scope of this chapter.

<sup>21</sup> I thank the participants in the 2019 Workshop in Normative Ethics and two anonymous referees for *Oxford Studies in Normative Ethics* for very helpful comments on this chapter. I also thank audiences at Cardiff, Oxford, and the 2013 Eastern APA (especially Kenneth Walden), as well as Joel Anderson, Tom Bates, Katharina Bauer, Jochen Bojanowski, David Copp, David Enoch, Govert den Hartogh, Florian Merwede, Hanno Sauer, Oliver Sensen, Paulien Snellen, Jens Timmermann, Marijana Vujosevic, Allen Wood, and members of the ethics research seminar at the University of Groningen for perceptive comments on a much earlier version of this chapter. I am grateful to the Dutch Research Council (NWO) for research support.

- Foot, Philippa. 1978 [1967]. "The Problem of Abortion and the Doctrine of Double Effect." In: *Virtues and Vices and Other Essays in Moral Philosophy*, 19–32. Oxford: Blackwell. [Originally in *Oxford Review* 5: 5–15.]
- Fried, Barbara. 2012. "What Does Matter? The Case for Killing the Trolley Problem (or Letting It Die)." *Philosophical Quarterly* 62: 505–29.
- Gigerenzer, Gerd. 2008. "Moral Intuition = Fast and Frugal Heuristics?" In: *Moral Psychology*, ed. Walter Sinnott-Armstrong, vol. 2, 1–26. Cambridge, MA: MIT Press.
- Greene, Joshua D., R. Brian Sommerville, Leigh E. Nystrom, John M. Darley, and Jonathan D. Cohen. 2001. "An fMRI Investigation of Emotional Engagement in Moral Judgment." *Science* 293: 2105–8.
- Hurka, Thomas. 2016. "Trolleys and Permissible Harm." In: F.M. Kamm, *The Trolley Problem Mysteries*, ed. Eric Rakowski, 135–50. Oxford: Oxford University Press.
- Kamm, F.M. 2007. *Intricate Ethics: Rights, Responsibilities and Permissible Harm*. Oxford: Oxford University Press.
- Kant, Immanuel. 1992–. *Cambridge Edition of the Works of Immanuel Kant*, eds. Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.
- Kant, Immanuel. 1900–. *Kant's gesammelte Schriften*, ed. Preussische [later Deutsche] Akademie der Wissenschaften. Berlin: Georg Reimer, later Walter de Gruyter.
- Kerstein, Samuel. 2013. *How to Treat Persons*. Oxford: Oxford University Press.
- Kleingeld, Pauline. 2020. "How to Use Someone 'Merely as a Means.'" *Kantian Review* 25: 389–414. DOI: 10.1017/S1369415420000229.
- Korsgaard, Christine M. 2009. *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Korsgaard, Christine M. 1996. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- O'Neill, Onora. 1989. *Constructions of Reason*. Cambridge: Cambridge University Press.
- Otsuka, Michael. 2008. "Double Effect, Triple Effect and the Trolley Problem: Squaring the Circle in Looping Cases." *Utilitas* 20: 92–110.
- Parfit, Derek. 2011. *On What Matters*. Oxford: Oxford University Press.
- Pettit, Philip. 2015. *The Robust Demands of the Good: Ethics with Attachment, Virtue, and Respect*. Oxford: Oxford University Press.



- Scanlon, T.M. 2008. *Moral Dimensions: Permissibility, Meaning, and Blame*. Cambridge, MA: Harvard University Press.
- Thomson, Judith Jarvis. 1976. "Killing, Letting Die, and the Trolley Problem." *Monist* 59: 204–17.
- Thomson, Judith Jarvis. 1985. "The Trolley Problem." *Yale Law Journal* 94: 1395–415.
- Thomson, Judith Jarvis. 2008. "Turning the Trolley." *Philosophy and Public Affairs* 36: 359–74.
- Thomson, Judith Jarvis. 2016. "Kamm on the Trolley Problems." In: F.M. Kamm, *The Trolley Problem Mysteries*, ed. Eric Rakowski, 113–34. Oxford: Oxford University Press.
- Unger, Peter. 1996. *Living High and Letting Die: Our Illusion of Innocence*. Oxford: Oxford University Press.
- Wood, Allen W. 2008. *Kantian Ethics*. Cambridge: Cambridge University Press.
- Wood, Allen W. 2011. "Humanity as an End in Itself." In Derek Parfit, *On What Matters*, vol. 2, 58–82. Oxford: Oxford University Press.