Philosophia (2014) 42:251–269
DOI 10.1007/s11406-013-9507-5

A Review and Systematization of the Trolley Problem

Stijn Bruers · Johan Braeckman

Received: 10 June 2013 / Revised: 11 October 2013 / Accepted: 28 October 2013 /
Published online: 12 November 2013

Abstract The trolley problem, first described by Foot (1967) and Thomson (*The Monist*, 59, 204–217, 1976), is one of the most famous and influential thought experiments in deontological ethics. The general story is that a runaway trolley is threatening the lives of five people. Doing nothing will result in the death of those persons, but acting in order to save those persons would unavoidably result in the death of another, sixth person. It appears that, depending on the situation, we have different moral judgments about the permissibility of action. We will review and systematize all the proposals in the literature of the past 35 years that have attempted to grasp our moral intuitions in a simple deontological principle. In particular, seventeen proposals will be classified: six algorithmic, seven psychological, and four other invalid accounts. This review and classification sheds light on some subtle differences and clarify a few issues.

Keywords Deontological ethics · Doctrine of double effect · Mere means principle · Risk aversion · Trolley dilemma

Introduction

The trolley problem consists of a series of moral dilemmas involving a runaway trolley threatening the lives of a certain number of people.[1] The basic structure of all the dilemmas is the same: if you do not act, five people will die; if you act, one other person will be killed and the five will be saved. Research into the way people deal ethically with the trolley dilemmas has shown that most people's intuitions do not correspond either with pure (extreme) deontology or with utilitarianism (Greene et al. 2001; Waldmann and Dieterich 2007; Hauser et al. 2008). By "pure" deontology we mean here, for simplicity's sake, that people should comply with the following rule: never act if the act results in harming people who were not threatened if you had not acted. By "pure"

[1]For a highly readable overview and historic background of 'trolleyology', with a focus on experimental (neuro)ethics, see Edmonds (2013).

S. Bruers (✉) · J. Braeckman
Department of Philosophy and Moral Science, Ghent University, Ghent, Belgium
e-mail: stijn.bruers@ugent.be

utilitarianism we mean that people should comply with the rule: always choose the action that maximizes the number of lives saved (i.e., least total harm).

Different trolley dilemmas have the same consequential structure but yet, confronted with those dilemmas, people hardly ever say that one should never act, or that one should always act. When presented with different dilemmas, most people say that we must act in one trolley situation, but in another dilemma we are not allowed to act; it is as if people make inconsistent choices. Only pure utilitarian consequentialism states that we should always act in all the trolley dilemmas. So most people's moral intuitions deviate from these consequentialist ethics, and therefore the trolley problem is an interesting thought experiment for studying deontological ethics. The basic question is the following: What is the morally relevant difference between Dilemma A and Dilemma B, such that it is morally allowable to act in A, but not to act in B? Also, in this article we state that a consistent moral solution of the trolley problem should contain a clear description of a rule or principle that best fits, justifies, and explains the diversity of people's moral intuitions in the diverse cases. In other words, the best solution to the trolley problem is a clear algorithm to decide whether one should act or not, and the answers that this algorithm generates should be in line with intuitions.

Many people have tried to solve the abovementioned basic question. In this article, we present an overview of the many proposals that ethicists have come up with during the last 35 years, and we discuss their differences, mutual relations, strengths, and weaknesses. In addition to covering the most relevant versions of trolley dilemmas and the solutions proposed in the literature, we also present some new hypothetical solutions. But probably the most important contribution of this article to the existing literature is a systematic classification of all those solutions. And looking at new trolley dilemmas, we clarify the differences between the proposed solutions (principles).

There are several reasons why this new classification is important. First, it certainly helps to avoid confusion between different solutions (we will mention some confusions in the literature). Second, the classification of the different principles gives us insights into which of those principles and underlying moral intuitions could be something like "moral illusions" (e.g., Unger 1996). Third, people adhering to deontological ethics might be able to see which proposed principle they would most prefer; that is, which of the proposals is most compatible with their own moral intuitions. Fourth, our findings will have implications for further empirical studies in moral psychology (e.g., Greene 2002; Cushman et al. 2006; Mikhail 2007; Greene 2008). This systematization opens up some new questions. Do people prefer one of the proposed solutions? How many people would agree with which solution? Would they change their judgments in some dilemmas in order to make them fit with their preferred solution? And if there are different proposed solutions related to different (psychological or algorithmic) mechanisms, does that mean that there would be more "moral modules" in our brains (e.g., the brain research on trolley dilemmas done by Greene et al. 2001)?

We start this review with a number of trolley dilemmas that cover all the important issues and elements that are discussed in the literature. Then, we select six possible solutions to the trolley problem that are described in the literature (the sixth is in fact a new solution), from which we will suggest that these can be grouped in pairs, so that there are in fact only three groups of principles with more or less strong support in the literature. These accounts have an algorithmic character, with a clearer and more objective decision procedure than the other, psychological accounts.

Of course, other hypothetical principles are possible, but they have few or no supporters or they remain dubious and are still debated. So after describing the three groups of "algorithmic" accounts, we give an overview of seven other proposals that do not distinguish permissibility from non-permissibility so clearly. Some of these other proposals might be relevant, as they are more "psychological" in nature and psychology strongly influences our moral judgments. However, these psychological explanations are not always clear or do not always make a distinction *between* the different "agent-neutral" trolley dilemmas: they make distinctions *within* one dilemma. An agent-neutral trolley dilemma is a description of a situation that excludes agent-related information.[2] The inclusion of agent-related information such as a person's position (e.g., distance from the victim) or mental state (e.g., knowledge, risk attitude, intention to harm) give rise to further distinctions within the same agent-neutral trolley dilemma.

After the six algorithmic and seven psychological accounts, we finally briefly highlight four other proposals encountered in the literature that do not solve the trolley problem, because they result in pure deontology. In summary, seventeen proposals in the literature are classified as follows: six solutions that make objective (algorithmic) distinctions between different dilemmas; seven solutions that make distinctions within one and the same dilemma, depending on some psychological state; and four invalid proposals that always result in pure deontology in all dilemmas.

So let's start the trolley's engine.

The Trolley Dilemmas

In this section we will briefly present and systematize the most commonly discussed versions of the trolley dilemma (for further details, see Thomson 1985; Kamm 1989, 1998; Otsuka 2008; Fischer and Ravizza 1992b).

> *Dilemma 1: The switch.* A trolley is moving towards five people on the main track. You are standing at a switch. If you turn the switch, the trolley will be diverted to a side track, but there is one person on this side track. Turning the switch will result in that person's death, and the five people on the main track will be saved. Should you turn the switch? Most people (roughly 90 % according to Hauser et al. 2008) say you are allowed to do so.
>
> *Dilemma 2: The bridge.* A fat man is standing on a bridge above the track. You can save the five people on the track below by pushing the fat man from the bridge in front of the trolley, so that the trolley will be stopped by his heavy weight. The fat man will die, and the five people will be saved. Only a few people (roughly 10 % according to Hauser et al. 2008) say that we are allowed to push the fat man. Most people either refuse to push the fat man or condemn pushing the fat man. According to Waldmann and Dieterich (2007), people are more tolerant of pushing someone onto the tracks, but in their study, the dilemmas did not involve close up and personal contact with the victim who is pushed. Their dilemmas looked more

[2] The only (trivial) agent related information in all situations is that if the agent acts, it is supposed that s/he acts with the intention or plan to save the people on the main track. I.e. malicious intentions (e.g. to kill a hated person) are excluded.

like Dilemma 5 below, where the victim is in a truck, so you do not have to touch the victim personally.

Dilemma 3: The loop. As in the first dilemma, you are standing at a switch. But this time the side track turns back onto the main track. If there is no one on the side track, the trolley will still move onto the main track and will kill the five people. But on the side track is a fat man. So if you turn the switch, the fat man will block the trolley. The fat man dies, the five people will be saved. In a recent survey, Hauser et al. (2008) found that roughly half of the respondents said that turning the switch is permitted. However, according to Waldmann and Dieterich (2007), people are more tolerant towards turning the switch. But they constructed the dilemma in a different way, where the person – the fat man – on the side track is sitting on a bus. So on the side track the bus will block the trolley, not the fat man. The fat man in the bus will die in the accident.

Dilemmas 1 to 3 share a rather similar basic structure, as can be seen in Fig. 1 below. In the first picture, the people on the main track are standing between two forks and the trolley is situated before the first fork. This is equivalent to Dilemma 1 (the switch). In the second picture, the five people are behind the second fork, and you have hesitated so long that the trolley already passed the first fork. You could still save the people on the main track, because the side track is at a height, so you could easily push the fat man from the side track onto the main track. In the third picture, the people on the main track are behind the second fork, as in Dilemma 3 (Loop), and you still have time to turn the switch. The difference is in the positions of the trolley and the people on the main track.

Dilemma 4: The loop with a stone. The situation resembles the one in Dilemma 3. But this time a heavy stone is located behind the man on the side track. The man on the side track is not heavy enough to stop the trolley. The stone will block the trolley if you turn the switch. But the man in front of the stone will die. In a survey

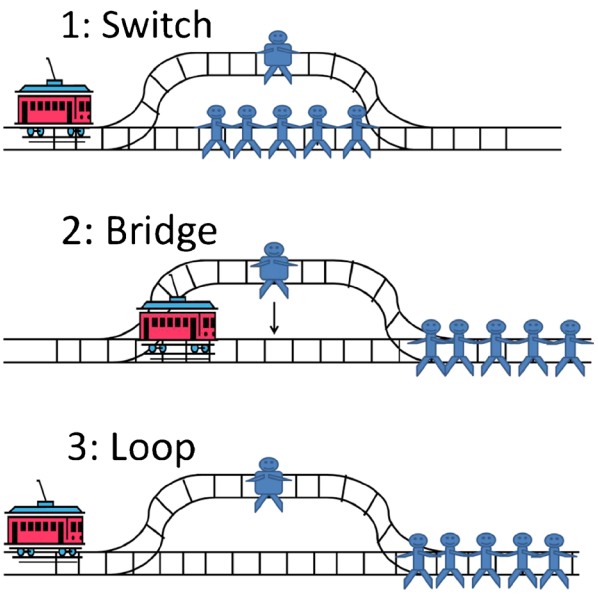

Fig. 1 Basic structures of Dilemma 1 (*Switch*); Dilemma 2 (*Bridge*); and Dilemma 3 (*Loop*)

performed by Hauser et al. (2008), it was found that nearly three quarters of the respondents say that turning the switch is allowed, and this is a statistically significant difference from Dilemma 3.

Dilemma 5: The truck. You can block the trolley by pushing a heavy truck onto the rails. In this truck there is one passenger, who will get killed (this dilemma was studied by Waldmann and Dieterich 2007).

Dilemma 6: The rockslide. You can turn a switch, redirecting the trolley onto a side track. On this side track there is a big rock. When the trolley hits the rock, the rock slides towards a bystander and kills him.

Dilemma 7: The platform. Five people are on a moving platform on the rails. If you do nothing, the trolley will crush the platform and kill the five. But you can move the platform away from the rails in order to save the five. But this move will push another person (who is standing next to the platform) on to an electric cable. This person will consequently die by electrocution. This dilemma is slightly similar to the "Lazy Susan" case in Kamm (1989).

The next dilemmas is new in the literature, and will be used to point out differences between some accounts.

Dilemma 8: The loop with an avalanche. This dilemma is similar to the loop dilemma, but the person on the side track is controlling a safety barrier against avalanches. The person on the side track is not heavy enough to stop the trolley, so one needs to create an avalanche. But only after the person dies is it possible to initiate an avalanche that is not blocked by the barrier; so only after the victim dies and is no longer controlling the barrier, can the trolley be blocked and the five saved.[3]

Six Algorithmic Accounts

Ethicists have looked for morally relevant differences between the above dilemmas. They want to find a moral rule that generates answers that are consistent with the answers (intuitive moral judgments) of the majority of people. As we have seen, the majority of people are very clear about the first two dilemmas, the switch and the bridge. Therefore, we think that a good solution should consist of a most precise formulation as possible of a moral criterion to distinguish the dilemmas, such that action is at least permissible in the switch dilemma but not in the bridge dilemma. Concerning the other dilemmas, there is less consensus about people's moral intuitions, so the solutions might differ in these cases. We prefer a moral rule that works like a kind of algorithm; that is, a clear procedure applicable to all dilemmas, which provides an unambiguous answer as to whether action is allowed or not, and without reference to fuzzy or ambiguous concepts.

There are five algorithmic accounts proposed in the literature. We will introduce a sixth. They can be grouped together in pairs; hence, we have structured them into three groups of accounts. The first proposal in each group is a rather vague account, vulnerable to misinterpretations or borderline cases. The second account in each group

[3] We note that Lippert-Rasmussen (1996) gave another dilemma that has a resemblance to Dilemma 8, according to the accounts mentioned in the next section.

is more precise, leaving less room for interpretation. In other words, the second accounts correspond to more accurate interpretations of the first ones in each group. We will apply these algorithmic accounts to the above dilemmas and show that these three explanations are different from each other.

Group A: The "Mere Means" Accounts

> *A1: Use as merely a means to an end*. The Kantian inspired right not to be treated solely as a means is based on the unalienable dignity of persons. This right trumps the right to life of other persons: it is never allowable to kill and use a person as merely a means – even if this means that by this act the lives of others could be spared. It is considered disrespectful to treat someone merely as means. This mere means account is mentioned in Thomson (1985), and as we will see in a later section, it is related to an interpretation of the doctrine of double effect (Quinn 1989b).

Looking at the dilemmas, we can say that the fat men on the bridge (Dilemma 2) and on the loop track (dilemma Loop) are used as "trolley blockers" or "human shields." So only in these dilemmas is action not allowed (they violate the dignity of the victim). In the other dilemmas action is permitted.

This account can sometimes be a bit vague, as it is not always easy to understand what use as merely a means really is. Sometimes one might have to use an element of fantasy to refer to an analogous means or instrument. We believe the following account is equivalent to this mere means account, but it provides a more algorithmic way (a clear test) to decide whether the Kantian right is violated or not.

> *A2: The counterfactual account about the required presence of the victim* (mentioned in Thomson 1985; Parfit 2011). If the presence of the potential victim is (causally) required in order to save the five people on the main track (i.e., if it would be impossible to save the five people without the victim's body), then you should not act. Here we can easily decide whether the Kantian right is violated, by asking ourselves what would happen if the one person in the trolley dilemmas was not present. If nobody is on the bridge, it is impossible to push someone in front of the trolley to stop it. But in considering Dilemma 1 (Switch), saving the five would still be possible if the one person on the side track had not been present. In Dilemma 4 (Loop and stone) one can still turn the switch and let the trolley be stopped by the stone. In Dilemma 5 (Truck) one can still move the truck, even if there is nobody in it. The platform (Dilemma 7) can still be shifted when there is no one standing next to it. In Dilemma 8 (Loop with avalanche) one could still start an avalanche when the victim on the side track is not present.

According to Waldmann and Dieterich (2007), using a person as a means is not a criterion that people use, because a lot of people say that we are allowed to turn the switch in Dilemma 3 (Loop). However, in their study, the person on the side track was sitting in a bus, and people might think it was the bus that is blocking the trolley. So the bus is used as the means, not the person's body. If the person was not sitting on the bus, saving the five by turning the switch still works. Furthermore, Hauser et al. (2008) noted a slight difference in people's responses between Dilemmas 3 (Loop) and 4 (Loop and stone). This difference can only be explained by the mere means accounts.

The mere means account also has another property: if the victim's body needs to be present (if there is no other heavy object that could replace the person as a trolley blocker), it also implies that there is logically no possibility of saving the one person after the five people on the main track are saved. After turning the switch in Dilemma 4 (Loop with stone), you could still try to save the person on the side track. But in Dilemma 3 (Loop), saving this person is impossible: even if you manage to run to the person and pull him away from the tracks, you cannot do this without endangering the five people again. This property might point at an evolutionary explanation of the moral intuition: in rescuing members from your group, it is advantageous to choose the option that allows you to try to save all of them. Saving everyone is not logically impossible in Dilemma 1 (Switch) and Dilemma 4 (Loop with stone). You first save the five people and then run to the side track. If you do not run fast enough, you still have saved the five, and if you can run fast enough, you can save everyone. The latter is logically impossible in Dilemmas 2 and 3 (Bridge and Loop).

Compared with the next four accounts, the mere means accounts are the most reliable: they are more accurate, have less boundary cases and generate less judgments that are strongly counter-intuitive in some dilemmas (the only counter-intuitive judgment occurs in the loop dilemma).

Group B: The "Same Threat" Accounts

The two same threat accounts that we are about to discuss have something in common; they both claim that it is only permissible to act if two conditions are satisfied: (1) no new threat is introduced, but a pre-existing threat is redirected or redistributed from the larger to the smaller group (this is also referred to as the Permissible Diversion Hypothesis in Postow 1989); and (2) another condition is satisfied – about this latter condition, we will discuss two candidates (related to rights or interventions), but we expect that they are equivalent.

The first condition is not satisfied in Dilemma 6 (Rockslide) and Dilemma 7 (Platform), because electrocution and rockslides are new threats. So in these dilemmas, action is already not permitted. However, a problem of this condition is its lack of clarity: it is not always clear when a new threat is introduced. There are some borderline cases of redirected threats that more resemble new threats (e.g., situations where trolleys change after being sent to a side track). These borderline cases might undermine the objective, algorithmic nature of the same threat account. Leaving this issue aside, let's look at the second condition. There are two versions of the second condition, leading to two accounts.

B1: Violation of rights. Thomson (1976, 1985) had the idea that action in the case of Dilemma 2 (Bridge) is not allowable, because pushing the fat man is an infringement of an important right. On the other hand, in Dilemma 1 (Switch) turning the switch does not violate a similar right of the person on the side track because you do not do anything to him. In particular, you do not push him, so his right not to be pushed is not violated. In other words in Dilemma 2 (Bridge), you do something to a person (which is a violation of rights), whereas in the switch you act on the threat (which is not a rights violation). Also in all the loop dilemmas (3, 4, 8) nobody is pushed, so no right not to be pushed is violated. Turning the switch

is itself morally neutral and not a violation of rights. Therefore, Thomson claimed that it is permitted to turn the switch in the loop dilemmas.

Thomson's idea has been criticized as being too vague and for contradicting moral intuitions (Kamm 1989; Postow 1989). What rights are we talking about – the right not to be killed by trolleys or the right not to be pushed? The following describes another candidate condition, which in fact might be equivalent to what Thomson had in mind, but stated a bit more clearly.

B2: Sending a victim to the trolley. Some ethicists claim that there is a morally relevant difference between throwing a bomb at a person and throwing a person at a bomb; or in trolley language, sending a trolley to a person versus sending a person to a trolley. This is referred to as intervention myopia (Waldmann and Dieterich 2007) and focuses at the locus of intervention: do you in the first instance intervene in the path of the aggressor (the trolley, the bomb) or the path of the victim? This criterion has some supporters (Boorse 1994; Harris 2000; Waldmann and Dieterich 2007) and some critics (Fischer 1992; Fischer and Ravizza 1994). Montmarquet (1982) also offered a same threat principle, but this approach was criticized by Kamm (1989).

Only in the bridge and truck dilemmas does one sends the victim (the fat man or the passenger) to the trolley. In these cases, action is not allowed, even if the threat is the same. We also note that in Dilemma 7 (Platform), the victim is sent to a new threat (electrocution), so both conditions of the same threat account are not satisfied. Yet, we expect that most people's intuitions would allow action in this dilemma. This gives a strong counter example to the same threat account.

Furthermore, there are some boundary cases between sending victims and threats to each other. As an example, consider a loop dilemma whereby turning the switch also shifts a platform, positioning the victim exactly on the side track to block the trolley. Also, according to Unger (1996, p. 101), the difference between sending the victim to the trolley versus sending the trolley to the victim is an illusion, based on what he called "protophysics."[4] In his book, Unger (1996) also gives other similar irrelevant protophysical differences that influence our moral judgments. For example: in some dilemmas it is worse to save some people and harm someone else by increasing the speed of a trolley than by decreasing it.

Group C: The "Causal Chain" Accounts

The next two principles look at the causal chain that is the result of action or inaction. We note that these principles should be taken with a grain of salt, because a clear and consistent interpretation of them might just be impossible if we think about them more critically. Nevertheless, we present them here.

[4] The loop dilemma is often used by some philosophers (e.g., Singer 2005; Scanlon 2008) to demonstrate the invalidity of the deontological mere means account, the abovementioned mere means principle, by claiming that a lot of people have the intuition that it is permissible to act in the loop dilemma, even when the victim is used as a trolley blocker. However, if this protophysical explanation is correct, the judgment in the loop dilemma (the permissibility to act), might be a moral illusion. We will demonstrate in more detail in another study.

C1: Principle of (Im)Permissible Harm (PI/PH). We cite Kamm (1989, in Darwall 2003, p. 167), who introduced this hypothesis: "It is permissible to cause harm to some in the course of achieving the greater good of saving a greater number of others from comparable harm, if events which produce the greater good are not more intimately causally related to the production of harm than they are to the production of the greater good." This is a complicated formulation that needs more explication, so let us look at the dilemmas to see what is meant by "intimately causally related."

Looking at the switch dilemma, the action is turning the switch, and this action has two consequences that appear at the same instant in the causal chain: the five are saved, and the one is threatened by the trolley. The production of the harm (the threat to the one on the side track) is causally related to the action of turning the switch, and also the saving of the five is causally related to the turning of the switch. Both are in this dilemma equally intimately causally related to the turning of the switch, because both are the direct consequences of turning the switch. The condition of the PI/PH is satisfied, so it is allowable to turn the switch.

In the bridge dilemma, however, the action is pushing the fat man. As a first consequence, the fat man is threatened; a second consequence is that the fat man blocks the trolley and saves the five. But looking at the causal chain, we see that the "causal distance" between the action (the pushing) and the harm (or the threat) to the fat man is smaller than the causal distance between the action and the saving of the five. The saving of the five happens further up in the causal chain. Therefore, PI/PH says that action is not permitted.

As it is not always clear how to calculate intimate causal relatedness, there are some borderline cases. In Dilemmas 3 (Loop), the situation is similar to Dilemma 1 (Switch), according to Kamm (1989).[5] However, we might disagree with this, as can be seen in Situations 2 and 3 (Fig. 1). In both the bridge and the loop dilemmas, the fat man is simply placed in the path of the trolley, either by changing the path of the trolley (Situation 3, Loop) or changing the position of the fat man (Situation 2, Bridge). Causally speaking, both are equivalent. So we should be a bit skeptical about this account.

The sixth possible explanation is not mentioned in the literature (as far as we are aware), and perhaps it is identical to an interpretation of Kamm's PI/PH hypothesis above. It is also vulnerable to borderline cases, such as the loop dilemma.

C2: The increased threat account. This rule says: you are allowed to act if the victim of your action dies (is harmed) after the others are saved. In the causal chain, we can represent it as follows. Each person has a value: 1 equals *alive and saved*, 0 means *dead (or harmed)*, X means *actually threatened (but still alive)*, and Y means *potentially threatened* (meaning it is possible to turn a threat towards that person). The six persons in the dilemmas all have a value at each step; so we can represent the starting situation as X,X,X,X,X,Y, that is, five persons are actually threatened and one person is potentially threatened (we are able to act so that his position would become a threatened one). Turning the switch in Dilemma 1 changes the situation to

[5] In her later work, Kamm (2007) introduced new refinements (e.g. causal versus non-causel flip sides, directly versus indirectly causing a lesser evil, producing versus sustaining a greater good and substituting versus subordinating persons). But this was criticized by Norcross (2008) as being heavily ad hoc and unclear.

Y,Y,Y,Y,Y,X, which means that one person is really threatened and the five people are potentially threatened: we can turn the switch back to change the situation back to the initial situation. So the five people are not yet absolutely sure about their survival. At a particular point, when the trolley passes the bifurcation in the track, the five people are actually saved and the situation turns into 1,1,1,1,1,X. And after a few moments, the person on the side track dies, resulting in 1,1,1,1,1,0. So the causal chain in the switch dilemma looks like:

$$X, X, X, X, X, Y \rightarrow Y, Y, Y, Y, Y, X \rightarrow 1, 1, 1, 1, 1, X \rightarrow 1, 1, 1, 1, 1, 0.$$

Action is allowable if the causal chain looks like that above. However, in Dilemma 2 (Bridge), matters are more complicated: it all depends on whether the fat man is heavy enough to block the trolley (i.e. whether the five are definitely saved once you pushed the fat man). It might be the case that the trolley is too fast and is able to kill all six people, because all six people are placed in the trajectory of the trolley. In other words, it is not clear that the five people are absolutely saved *already at the moment when one pushes the fat man*. The causal chain now can look like:

$$X, X, X, X, X, Y \rightarrow X, X, X, X, X, X \rightarrow 1, 1, 1, 1, 1, 0.$$

The possibility of the X,X,X,X,X,X situation (everyone is in danger), distinguishes the bridge from the switch dilemma. The number of threatened people is increased. In Dilemma 8 (Loop with avalanche) we clearly see a moment where everyone is in danger: when turning the switch and the trolley passes the bifurcation, the one person on the side track is threatened, but the five other people are also still threatened, because the one person is not heavy enough to block the trolley. Only after the one person is killed does it become possible to relieve the threat to the five, by initiating an avalanche that blocks the trolley. The causal chain in Dilemma 8 looks like:

$$X, X, X, X, X, Y \rightarrow X, X, X, X, X, X \rightarrow X, X, X, X, X, 0 \rightarrow 1, 1, 1, 1, 1, 0$$

The causal chain account is only able to distinguish Dilemma 1 (Switch) from Dilemma 2 (Bridge) if the causal chain can have a point where everyone is in danger, for example if we suppose that the train might kill all six people in Bridge. The existence of this increased threat situation in the causal chain disallows action. But then we have to suppose a similar possibility in Dilemma 3 (Loop), which disallows action. However, Dilemma 4 (Loop and stone) becomes creates a boundary case: if the stone was a real mountain, the possibility of situation X,X,X,X,X,X is as unlikely as it is in Dilemma 1 (Switch). So the permissibility depends on whether the stone is really heavy enough to block the trolley. But the same can be said about the fat man, who might be heavy enough to block the trolley. If we know the fat man is heavy enough, there is no distinction between Switch and Bridge: after pushing the fat man, the path of the trolley changes (it stops), just as the path of the trolley changes in the switch dilemma. Pushing the fat man and turning the switch automatically guarantee the immediate safety of the five people. We can introduce a distinction by claiming that the agent cannot be sure

whether the fat man is heavy enough, but this turns the algorithmic account into a psychological account (see the section in risk aversion below).

The increased threat account reveals a kind of "causal myopia" (similar to the term "intervention myopia" related to the same threat account; Waldmann and Dieterich 2007). If in the series of consequences of your action you do not threaten someone before or at the moment when others are really saved, then you are allowed to act. It is as if you were blind to the further consequences in the causal chain.

Table 1 presents the results of the trolley dilemmas according to the above three principles: the mere means, same threat, and causal chain accounts. A plus means that the action is allowed, a minus means that it is not. As discussed, the causal chain accounts have some question marks.

Note that the switch and bridge cases get all plus and minus signs respectively, so for these all three accounts can be considered as a solution to the trolley problem. But the answers differ when looking at other dilemmas.

Seven Psychological Accounts

The accounts that we presented above are all objective, in the sense that they did not refer to mental states, but to events, counterfactual requirements, number of threats, points of intervention, directions or causal consequences. In this section, we give an overview of some other proposals discussed in the literature. These proposals often involve some psychological influences, such as intentions, risk aversion, personal versus non-personal conflict, and so on. These psychological accounts have some flaws: sometimes they are not able to derive a clear judgment in a certain dilemma (especially the loop dilemma generates problems of interpretation), or they do not always make a clear distinction *between* dilemmas such as the switch and the bridge. They make distinctions even *within* one dilemma. So depending on the situation (related to the psychological states) it might be possible that it is not permitted to act in the switch dilemma, or that it is permitted to act in the bridge dilemma.

1. *The Doctrine of Double Effect (DDE).* This doctrine is mentioned in quite a few discussions about the trolley problem (Boyle 1980; Davis 1984; Fischer and Ravizza 1992a; Reibetanz 1998; McIntyre 2001; Shaw 2006; Edmonds 2013).

Table 1 Answers to the trolley dilemmas, according to the three accounts

Dilemma	Mere means account	Same threat account	Causal chain account
1. Switch	+	+	+
2. Bridge	−	−	−(?)
3. Loop	−	+	−(?)
4. Loop and stone	+	+	−(?)
5. Truck	+	−	−(?)
6. Rockslide	+	−	+
7. Platform	+	−	+
8. Loop and avalanche	+	+	−

The doctrine says that there is a moral difference between the intentional harm as a means and the foreseen harm as a side-effect (Quinn 1989b). It has been criticized by, for example, McIntyre (2001).

The DDE is an agent-centered, psychological account, as it makes a difference between what the agent intends or foresees.[6] We could try to interpret the DDE in a more agent-neutral way; that is, without too much reference to the mental states of agents. Reinterpreting the DDE as an agent-neutral principle moves it close to the mere means account discussed above, because the DDE refers to "harm as a means." However, we have to be careful not to confuse the use of a person's body as a means versus the use of, for instance, a switch as a means or a plan as a means.

The difference between the DDE and the mere means accounts can be most clearly seen in Dilemma 8 (Loop and avalanche). The person on the side track is not used as a means, because the presence of his body is not necessary to save the five (on the contrary, his presence has prevented the initiation of the necessary avalanche). But the agent intends the killing (removal) of the person on the side track, because this removal is necessary in order to initiate the avalanche. The DDE says that action is not allowed, because it involves an intentional harm.

Hence, the DDE is not simply equivalent to the mere means account, a fact that might result in misinterpretations in the literature. For example Costa (1986), in his application of the DDE to the trolley dilemma, combined (or confused?) the mere means account with a version of the causal chain account. And to make it even more extraordinary (or confusing), in a later article Costa (1987) also included a version of Thomson's "same threat" principle in the description of the DDE, as if the DDE is a confusing mixture of all three groups of accounts discussed in the previous section.

The major problem with the DDE is the loop dilemma: is the death of the person on the side track intended or merely foreseen? When Kamm (2000) tried to apply the DDE to the loop trolley dilemma, she promoted a new doctrine of triple effect (DTE).

2. *The Doctrine of Triple Effect (DTE).* Following Kamm's doctrine, turning the switch in the loop case is permissible according to triple effect. That is because apart from intentional harm (doing something *in order to* bring about an evil) and merely foreseeing a side effect (doing something *in spite of* bringing about an evil), Kamm claims that there is a third option, in which one does something *because* it brings about an evil (which should be distinguished from "in order to bring about an evil"). This DTE approach was further defended by Shaw (2006) but criticized by Harris (2000) and more recently by Otsuka (2008) and Liao (2009) using the loop dilemma: triple effect does not solve the loop case either. Liao argued that the because of/in order to distinction does not apply to the loop case, and furthermore questions whether this distinction has a normative significance.

Otsuka (2008) gives an example of a trolley dilemma where this triple effect becomes clearer: suppose you are at a switch, and on the side track there is one person in front of six other people. If you turn the switch, the five on the main track are saved, the first person on the side track will block the trolley, and the six people

[6] Of course, the objective accounts also include a trivial mental state of the agent: if the agent acts, s/he is supposed to have an intention or plan to save the people on the main track. However, the DDE refers to a non-trivial mental state: the intention to harm (distinguished from foreseeing the harm).

behind him are saved. Here, we can say that we would turn the switch, not *in order to* kill the one on the side track, but rather *because* he will be killed and stop the trolley. Nothing new is added however, we think, because action in this dilemma is also allowed according to our abovementioned three accounts.[7]

3. *Feelings of the victim*. Thomson (1993) invited us to focus on what the potential victim would feel about what the agent does. If you were thrown from a bridge you might feel differently about the agent, than if a trolley were directed towards you. And it is this difference that plays a role. However, this claim also involves some complex knowledge of psychology, this time not of the agent, but of the victim. It does not yet solve the trolley problem, because one can imagine switch and bridge situations where the victim feels the same.
4. *Projective grouping*. Peter Unger speculated about another psychological mechanism behind our moral judgments: projective grouping and projective separating (Unger 1996, p. 97). "[When certain people are in a situation that is taken to be their problem], we tend to think it is badly wrong to spare them the serious losses that might stem from their problem by imposing serious loss on other people, who don't have that problem." In the first trolley dilemma (Switch), all six people on the tracks are considered to be in a similar position in that they have something in common: they are all on a track and could be run over by a trolley. So the five on the main track and the one on the side track are grouped together as having the same problem, and the one on the side track can therefore be considered as "fair game" to be sacrificed. However, in the bridge dilemma, the fat man is in a different position: he is not on a track, but on a bridge. So the fat man is psychologically separated from the five people on the track, which makes us decide not to sacrifice the fat man. A lot of people, when responding to the trolley dilemmas, give spontaneous answers that reflect this projective separation (people say something like, "But the fat man had nothing to do with it, he was just passing by"). Also Hanna (1992) proposed a Principle of Moral Inertia, which is basically the same as the projective separating. A distinction is made between participants (such as the person on the side track in the switch dilemma) who are part of an ongoing causal process, and bystanders (such as the fat man on the bridge) who are not part of the ongoing process. But this explanation is not fully satisfactory, however, because as Unger himself argued, it can be twisted. And it is at the least very vague: there are no clear criteria to separate people into groups. There is no consensus about what the relevant differences should be. Knowing whether someone is a participant or a bystander is not straightforward. And what about Situation 2 in Fig. 1, where the fat man was on a side track on a bridge?
5. *Epistemic accounts: risk aversion*. Risk aversion is a psychological attitude that might give an interesting explanation for the moral intuitions in the trolley problems. Can we know whether our plan to save the five would really work? If the fat man is not heavy enough and the trolley were to keep on moving, then all six will die, which is an even worse outcome. There is the risk of a worse outcome. If the trolley could have stopped in time, even without the fat man blocking it, then the

[7] For some further subtlety, however, we can say that the one person on the side track is a means to save the six behind him, but he is not used as means to save the five. If the person was not present, the plan to turn the switch and save the five would still work (but six other people would be threatened).

fat man would have died unnecessarily. In the switch dilemma, however, we can be pretty sure that the five are saved and nobody dies in vain.

According to this epistemic account, action would be impermissible if there is a possibility that the rescue plan will fail and all six people will die. In particular, action might not be allowed in Dilemmas 2, 3, and 5 (Bridge, Loop and Truck).

The problem with this hypothesis is that certainty is a matter of degree. Take Dilemma 4 (Loop and stone): What if the stone was really heavy so that you could be sure that it would stop the train? Surely a mountain of stones would be convincing. And even in the switch dilemma, suppose that the side track bends behind a hill. You cannot be sure that there are no people on the side track behind the hill. Perhaps there are ten people on the side track, but you cannot see them.

So the epistemic account in fact makes distinctions even within one dilemma, instead of between dilemmas. Nevertheless, there might be some interesting truth in this approach. It is related to the amount of risk aversion that the agent has. Suppose in the bridge dilemma there is a 10 % probability that the plan of pushing the fat man fails and all six people die instead of one, an 80 % probability that the plan will work and one person will die instead of five, and a 10 % probability that the trolley could have stopped anyway without the fat man, so that one person dies instead of nobody. A person with a high level of risk aversion would choose not to act. A person with maximum risk aversion would never act, even if the probability of failure were 0.0001 %. In this context, we note that most people have a high but not maximum level of risk aversion.

Going back to the switch dilemma, risk aversion would imply not turning the switch if there is a possibility that there are ten more people down the side track. But be aware that the same could apply to the main track: it might *equally* be possible that there are ten people behind the five, and you did not see them. Not turning the switch would result in fifteen deaths. Notice the word "equally". There is a kind of symmetry in the switch dilemma; whereas in most bridge dilemma situations that we imagine we do not see such a symmetry, and risk aversion has a stronger influence in those dilemmas.

6. *Epistemic accounts: uncertainty aversion.* Next to risk aversion there is uncertainty (or ambiguity) aversion, whereby the probabilities of success are not even known. The probability that the plan involving pushing the fat man will work is not 10 %: it is usually not known. So we have to choose between two games of chance. Suppose that you are one of the six people in the bridge trolley dilemma, but you do not know which one. If the fat man is not pushed, you know that the trolley will continue moving and kill five people. So you have a survival probability of one sixth, because you have a one sixth probability of being the fat man who survives. This is the first game of chance. In the second game, the fat man is pushed, and there is still a possibility that the trolley continues on and kills one or more of the five people on the track. Perhaps all might die. Which game of chance would you prefer to play? The situation is very similar to Ellsberg's paradox (Ellsberg 1961). Suppose we have an urn and you know three things: it contains six balls, has six (or fewer) different colors, and there is one green ball. The choice is between two games of chance. In the first, you win when you draw the green ball. Your probability of winning is one sixth. In the second, you win when you draw a blue ball. Your probability of winning is now unknown (somewhere between zero and five sixths),

because you do not know how many blue balls there are. Some (or most) people prefer the first game, because they have uncertainty aversion. The similarity with the trolley game is obvious.

7. *Personal* versus *impersonal dilemmas.* Greene (2008), finally, points – using psychological and brain research – to an important aspect in the trolley dilemmas: the distinction between personal versus impersonal dilemmas, related to the relative position of the agent towards the victim. Pushing the fat man is an action, which is close up and personal, whereas turning the switch is a more detached action. This is certainly something that influences people's choices, but it is not sufficient to solve the trolley problem, because it is easy to invent scenarios such as the bridge dilemma to make the action more detached (e.g., you are standing far away from the bridge and the fat man, but you can push a button, overturning the bridge). So this criterion would also make a distinction within the bridge dilemma. When most people imagine the bridge dilemma as a close up and personal situation, some emotion reaction in their brains will be triggered and tell them not to push the fat man.

   Interestingly, in their research, Greene et al. (2001) classified personal dilemmas using some criteria, one of them reads: "where this harm is not the result of deflecting an existing threat onto a different party" (Greene 2002, p168). This refers to the same threat account.

Four Invalid Accounts

In this section, we summarize some proposals that in fact would all result in "pure" deontology, so they do not solve the trolley problem.

1. *The Doctrine of Doing* versus *Allowing (DDA).* This principle of DDA is that there is a moral difference between killing and letting die. Quinn (1989a) for example, referred to the DDA to distinguish between trolley dilemmas. But as Fischer and Ravizza (1992a) argued, matters get very complicated in applying the DDA to trolley dilemmas, because one needs to include unsatisfactory references to concepts such as "transfer of intentions," "causal isolation," and so on.

   In line with the DDA, Foot (1978) made a distinction between positive duties (aid) versus negative duties (non-interference), and applied this to the trolley dilemmas. But this approach was criticized by Thomson (2008), who showed that Foot's idea basically results in pure deontology, whereby no action is permitted in all the trolley dilemmas.

   Interestingly, Thomson (2008) also ends up with pure deontology. However, we think that she is mistaken at some point. Thomson (2008, p.365) used a wrong argument (wrong analogy) to demonstrate that action is never allowed. Let us digress on this a little, because it is a recent discussion. Thomson starts with the "three options" dilemma: you are at a switch – if you do nothing the trolley will kill five people on the main track. You can also turn the switch to the right hand track, where one person will be killed, or turn it to the left hand track, where you will be killed. The argument goes that nobody is willing to sacrifice himself/herself (apart from real altruists or depressed people), and it is really unfair, Thomson claims, to

turn the switch to the one victim on the right hand track. To show that this is unfair, Thomson uses another example: You are asked to give money to a charity, in order to save people. You are able to send your own money, but you instead feel like stealing the money of someone else and sending that money to the charity. We claim that the analogy does not apply, because in the charity dilemma, you are *using* something of someone else. It is comparable to the "transplant dilemma" (Thomson 1985), whereby a surgeon can save five patients by sacrificing an innocent person and use that person's organs for transplantation. The transplant dilemma is similar to the bridge trolley dilemma, whereby you also use something of the victim, namely his body, without his consent. So Thomson's analogy can be used to argue that pushing the fat man in the bridge dilemma is not allowed. But from the charity dilemma analogy, it does not yet follow that turning the switch is not allowed. FitzPatrick (2009) and Shaver (2011) also commented on Thomson's new turn towards pure deontology (Thomson 2008).

2. *Illigitimate plans*. Russell (1979) referred to "illegitimate plans" to argue against the permission of, for example, pushing the fat man from the bridge. But this idea was criticized by Kamm (1989), and is in fact equal to pure deontology, as all the actions are shown to be illegitimate plans.
3. *Threatened persons.* Montmarquet (1982) stated that only when a person was not threatened is action impermissible. But he claimed that the person on the side track is already threatened. This claim, however, is false, as was argued by Gorr (1990). Montmarquet's approach would result in pure deontology, just like the DDA (Gorr himself, by the way, refers to a "same threat" account).
4. *Rare situations.* Gert (1993) claimed that – in contrast with impermissible actions – the permissible actions occur in very rare situations. Sitting in a truck next to a railway (Dilemma 5) or standing on a bridge above a railway (Dilemma 2) are more typical situations than standing on the railway with no escape possible, standing next to a moving platform, or standing in the pathway of a rock, and so on. The problem with this approach is that it is difficult to quantify the rarity of a situation and derive from this the permissibility of actions.

Conclusion and Further Research

The basic question we now have to ask is whether our pattern of answers (following our moral intuitions) is given by one or more of the three accounts presented above: the mere means, same threat, or causal chain accounts. If not, there must be other principles, or we must have moral illusions (comparable to optical illusions and cognitive biases). We know that there are some very peculiar examples of irrationalities in people's answers to the trolley dilemmas. Unger (1996) demonstrated that people's responses to the trolley dilemmas are often more (inconsistent) psychology than ethics, by pointing out that judgments about the permissibility of an option (e.g. the choice to push the fat man) depends on the availability of other options that people consider as being irrelevant (see also Norcross 2008). Furthermore, there is the well known effect of wording and framing (Petrinovich and O'Neill 1996; Sinnott-Armstrong 2008; Lanteri et al. 2008; Rai and Holyoak 2010). Especially the order in which different trolley dilemmas are presented, has some influence (Petrinovich and O'Neill 1996; Liao et al. 2011;

Schwitzgebel and Cushman 2012; Di Nucci 2012). And also induced feelings of disgust (Schnall et al. 2008) and happiness (Valdesolo and DeSteno 2006) can influence moral intuitions in the trolley dilemmas. The turn that the trolley problem has made towards empirical studies in moral psychology (Greene 2008; Cushman et al. 2006; Mikhail 2007) is very fruitful, especially in discovering moral illusions. Still, the abovementioned studies in experimental philosophy do not indicate that the gap between the two paradigmatic cases Switch and Bridge can be closed.

Given the classification of different accounts above, we can now ask the following (empirical) questions: How many people can agree with one or more of the three principles? Which account will have the most followers? What happens if respondents learn about these accounts? Do people feel satisfied with these accounts, and would they pick a preferred one? Will this influence their judgments in some dilemmas, and how? If, for example, a person is permissive towards the action in the loop dilemma, but learned that the mere means account is perfectly compatible with all of his/her intuitions in all dilemmas, except for the loop dilemma, would that change the judgment in the loop dilemma? Can such reflections easily override intuitions? Would that eventually influence the intuition in that dilemma? Can those accounts be used in the method of reflective equilibrium (Rawls 1971)? These can be questions for future research.

References

Boorse, C. (1994). Ducking trolleys. *Journal of Social Philosophy, 25*(3), 146–152.

Boyle, J. M. (1980). Toward understanding the principle of double effect. *Ethics, 90*(4), 527–538.

Costa, M. J. (1986). The trolley problem revisited. *The Southern Journal of Philosophy, 24*(4), 437–449.

Costa, M. J. (1987). Another trip on the trolley. *The Southern Journal of Philosophy, 25*(4), 461–466.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: testing three principles of harm. *Psychological Science, 17*(12), 1082–1089.

Darwall, S. (Ed.). (2003). *Deontology*. Oxford: Blackwell Publishing.

Davis, N. (1984). The doctrine of double effect: problems of interpretation. *Pacific Philosophical Quarterly, 65*, 107–123.

Di Nucci, E. (2012). Self-sacrifice and the trolley problem. *Philosophical Psychology,* Forthcoming.

Edmonds, D. (2013). *Would you kill the fat man? The trolley problem and what your answer tells us about right and wrong.* Princeton University Press.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics, 75*(4), 643–669.

Fischer, J. M. (1992). Thoughts on the trolley problem. In J. M. Fischer & M. Ravizza (Eds.), *Ethics: Problems and principles*. New York: Holt, Rinehart, & Winston.

Fischer, J. M., & Ravizza, M. (1992a). Quinn on doing and allowing. *The Philosophical Review, 101*(2), 343–352.

Fischer, J. M., & Ravizza, M. (1992b). *Ethics: Problems and principles*. New York: Holt, Rinehart & Winston.

Fischer, J. M., & Ravizza, M. (1994). Ducking harm and sacrificing others. *Journal of Social Philosophy, 25*(3), 135–145.

FitzPatrick, W. J. (2009). Thomson's turnabout on the trolley. *Analysis, 69*(4), 636–643.

Foot, P. (1967, 1978). The problem of abortion and the doctrine of double effect. In *Virtues and vices*. Oxford: Basil Blackwell, pp.19–32 (originally appeared in the *Oxford Review* 5, 1967.)

Gert, B. (1993). Transplants and trolleys. *Philosophy and Phenomenological Research, LIII*(1), 173–179.

Gorr, M. (1990). Thomson and the trolley problem. *Philosophical Studies, 59*(1), 91–100.

Greene, J. D. (2002). *The terrible, horrible, no good, very bad truth about morality and what to do about it.* Dissertation. Department of Philosophy, Princeton University.

Greene, J. D. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology, vol. 3: The neuroscience of morality*. Cambridge: MIT Press.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*, 2105–2108.

Hanna, R. (1992). Morality De Re: reflections on the trolley problem. In J. M. Fischer & M. Ravizza (Eds.), *Ethics: Problems and principles*. New York: Holt, Rinehart & Winston.
Harris, J. (2000). The moral difference between throwing a trolley at a person and throwing a person at a trolley: a reply to Kamm. *Proceedings of the Aristotelian Society, Supplementary Volumes, 74*, 41–57.
Hauser, M., Young, L., & Cushman, F. (2008). Reviving Rawls' linguistic analogy: operative principles and the causal structure of moral actions. In W. Sinnott-Armstrong (Ed.), *Moral psychology and biology*. NY: Oxford U. Press.
Kamm, F. M. (1989). Harming some to save others. *Philosophical Studies, 57*, 227–260.
Kamm, F. M. (1998). *Morality, mortality: Death and whom to save from it*. New York: Oxford University Press.
Kamm, F. M. (2000). The doctrine of triple effect and why a rational agent need not intend the means to his end. *Proceedings of the Aristotelian Society, Supplementary Volumes, 74*, 21–39.
Kamm, F. M. (2007). *Intricate ethics. Rights, responsibilities, and permissible harm*. Oxford: Oxford University Press.
Lanteri, A., Chelini, C., & Rizzello, S. (2008). An experimental investigation of emotions and reasoning in the trolley problem. *Journal of Business Ethics, 83*, 789–804.
Liao, M. (2009). The loop case and Kamm's doctrine of triple effect. *Philosophical Studies, 146*(2), 223–231.
Liao, M., Wiegmann, A., Alexander, J., & Vong, G. (2011). Putting the trolley in order: experimental philosophy and the loop case. *Philosophical Psychology, 25*(5), 1–11.
Lippert-Rasmussen, K. (1996). Moral status and the impermissibility of minimizing violations. *Philosophy and Public Affairs, 25*(4), 333–351.
McIntyre, A. (2001). Doing away with double effect. *Ethics, 111*(2), 219–255.
Mikhail, J. (2007). Universal moral grammar: theory, evidence, and the future. *Trends in Cognitive Sciences, 11*, 143–152.
Montmarquet, J. (1982). *Doing good: The right and the wrong way* (pp. 439–455). LXXIX: Journal of Philosophy.
Norcross, A. (2008). Off her trolley? Frances Kamm and the metaphysics of morality. *Utilitas, 20*(1), 65–80.
Otsuka, M. (2008). Double effect, triple effect and the trolley problem: squaring the circle in looping cases. *Utilitas, 20*(1), 92–110.
Parfit, D. (2011). *On what matters*. Oxford University Press.
Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology, 17*, 145–171.
Postow, B. C. (1989). Thomson and the trolley problem. *Southern Journal of Philosophy, 27*(4), 529–537.
Quinn, W. (1989a). Actions, intentions, and consequences: the doctrine of doing and allowing. *Philosophical Review, 98*, 287–312.
Quinn, W. (1989b). Actions, intentions, and consequences: the doctrine of double effect. *Philosophy and Public Affairs, 18*(4), 334–351.
Rai, T., & Holyoak, K. (2010). Moral principles or consumer preferences? Alternative framings of the trolley problem. *Cognitive Science, 34*, 311–321.
Rawls, J. (1971). *A theory of justice*. Cambridge: Harvard University Press.
Reibetanz, S. (1998). A problem for the doctrine of double effect. *Proceedings of the Aristotelian Society, 98*, 217–223. New Series.
Russell, B. (1979). The relative strictness of positive and negative duties. In B. Steinbock (Ed.), *Killing and letting die*. NY: Prentice-Hall.
Scanlon, T. M. (2008). *Moral d imensions*. Harvard University Press.
Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *PSPB, 34*(8), 1096–1109.
Schwitzgebel, E., & Cushman, F. (2012). Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. *Mind & Language, 27*, 135–153.
Shaver, R. (2011). Thomson's trolley switch. *Journal of Ethics & Social Philosophy*, discussion note.
Shaw, J. (2006). Intentions and trolleys. *Philosophical Quarterly, 56*(222), 63–83.
Singer, P. (2005). Ethics and intuitions. *The Journal of Ethics, 9*, 331–352.
Sinnott-Armstrong, W. (2008). Framing moral intuitions. In W. Sinnott-Armstrong (Ed.), *Moral psychology, volume 2: The cognitive science of morality* (pp. 47–76). Cambridge: MIT Press.
Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist, 59*, 204–217.
Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal, 94*, 1395–1415.
Thomson, J. J. (1993). Reply to commentators. *Philosophy and Phenomenological Research, 53*(1), 187–194.
Thomson, J. J. (2008). Turning the trolley. *Philosophy and Public Affairs, 36*(4), 359–374.

Unger, P. (1996). *Living high and letting die*. Oxford: Oxford University Press.
Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science, 17*, 476–477.
Waldmann, M., & Dieterich, J. (2007). Throwing a bomb on a person versus throwing a person on a bomb. Intervention myopia in moral intuitions. *Psychological Science, 18*(3), 247–253.