

What Happened to the Trolley Problem?

COVA, Florian

Abstract

In this paper, I provide a general introduction to the trolley problem. I describe its birth as a philosophical thought experiment, then its successful career in moral psychology. I explain the different reasons behind its popularity and success but argue that, despite its popularity and widespread utilization in psychological research, few researchers have actually tried to directly solve it and that we are still ignorant of the real factors guiding our responses to trolley cases. Against the idea that trolley problems were just a fad, I insist on the importance of seriously studying them and put forward a new hypothesis about the psychological mechanisms underlying our responses to trolley cases.

Reference

COVA, Florian. What Happened to the Trolley Problem? *Journal of Indian Council of Philosophical Research*, 2017, vol. 34, no. 3, p. 543-564

DOI : 10.1007/s40961-017-0114-x

Available at:

<http://archive-ouverte.unige.ch/unige:107299>

Disclaimer: layout of this document may differ from the published version.



UNIVERSITÉ
DE GENÈVE

What Happened to the Trolley Problem?

Florian Cova¹

Received: 3 May 2017 / Revised: 8 June 2017 / Accepted: 11 June 2017 /
Published online: 24 August 2017
© ICPR 2017

Abstract In this paper, I provide a general introduction to the trolley problem. I describe its birth as a philosophical thought experiment, then its successful career in moral psychology. I explain the different reasons behind its popularity and success but argue that, despite its popularity and widespread utilization in psychological research, few researchers have actually tried to directly solve it and that we are still ignorant of the real factors guiding our responses to trolley cases. Against the idea that trolley problems were just a fad, I insist on the importance of seriously studying them and put forward a new hypothesis about the psychological mechanisms underlying our responses to trolley cases.

Keywords Experimental philosophy · Ethics · Moral dilemmas · Moral psychology · Trolley Problem

Introduction: The Intuitive Turn in Moral Psychology

In the past decade, moral psychology, the empirical study of the psychological bases of moral judgment and moral behavior, has been blooming. According to a narrative widely shared within the discipline itself, this growth is due to a recent change in paradigm, which occurred around the beginning of the twenty-first century.

Here is how the story goes: In the second half of the twentieth century, moral psychology was mainly a branch of developmental psychology, and its focus was on the emergence of moral concepts in children (Turiel 1983) and the development of moral judgment and moral reasoning throughout the lifespan (Kohlberg and Hersh 1977). Because most of this research considered moral judgment as the result of

✉ Florian Cova
florian.cova@gmail.com

¹ Swiss Center for Affective Sciences, University of Geneva, Geneva, Switzerland

some form of reasoning, the study of the development of moral judgment was equated with the study of the development of moral reasoning.

This is why the method that dominated research in moral psychology at the time was a method developed by Lawrence Kohlberg, a former student of Piaget. The method consisted in having participants read small scenarios (typically moral dilemmas), form a moral judgment about them and then justify their moral judgment. Justifications provided by participants were taken to reflect the reasoning that drove and motivated their judgment in the first place. Thus, participants' skills (or 'levels') in moral reasoning could be evaluated from the kind of *arguments* they gave in favor of their judgment. The *content* of the judgment itself, however, was considered irrelevant to the assessment of the participants' moral development.

Among the scenarios used by Kohlberg and those who adopted his method, one came to gain popularity and can be seen as a 'flagship' for this method. Here is how it goes:

HEINZ STEALS THE DRUG – In Europe, a woman was near death from a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium that a druggist in the same town had recently discovered. The drug was expensive to make, but the druggist was charging ten times what the drug cost him to make. He paid \$200 for the radium and charged \$2000 for a small dose of the drug. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about \$1000 which is half of what it cost. He told the druggist that his wife was dying and asked him to sell it cheaper or let him pay later. But the druggist said: "No, I discovered the drug and I'm going to make money from it." So, Heinz got desperate and broke into the man's store to steal the drug-for his wife. Should the husband have done that? (Kohlberg 1963, p. 19)

What you think the husband should do is actually irrelevant. What matters is how you justify your answer. If you reason purely in terms of self-interest (e.g., the husband should steal the drug because he would be sad if his wife died *or* the husband should not steal the drug because he would be punished if he did), then congratulations! You are at the lowest level possible of moral reasoning (the *pre-conventional* level). If you reason in terms of what is generally approved in your society or in terms of what is authorized by the law of your country (e.g., the husband should steal the drug because this is what everybody expects from a loving husband *or* the husband should not steal the drug because theft is forbidden by the law), then you are at the intermediate level (the *conventional* level). Finally, if you reason in terms of universal moral principles or values (e.g., the husband should steal the drug because life is more important than property *or* the husband should not steal the drug because this would violate the druggist's property rights), then you have reached the highest, ultimate level of moral reasoning (the *post-conventional* level).

By the end of the twentieth century, Kohlberg's empirical paradigm, though influential, became the target of heavy criticisms. The most important criticism targeted Kohlberg's assumption that participants' justifications actually *reflect* the psychological processes that drove their moral judgment in the first place. Indeed,

this assumption commits Kohlberg to the view that we have a good introspective access to the cognitive processes that underlie our judgments or decisions. However, at the time, an increasing amount of empirical data suggested that such was not the case.

Indeed, in an influential review paper published in 1977, the psychologists Nisbett and Wilson famously argued that people do not have a full introspective access to the processes guiding their decisions and judgments and that their justifications for those often miss the true determinants of their behavior. Nisbett and Wilson grounded their claim on the results of a series of astonishing experiments. In one of these experiments, participants were instructed to choose one of several pairs of pantyhose displayed in a row. All items on display were perfectly identical. Still, participants displayed a preference for items on the right-hand side of the display. This preference could not be explained by some unique property possessed by the items on the right-hand side—rather, it was only due to a widespread bias to favor items on right, all things being equal. However, participants were not aware of this bias: Rather, when asked to justify their decision, participants tried to give rational justifications for it, mentioning reasons such as superior knit, sheerness and elasticity.

These results, as well as countless others, suggest that the justifications we give for our behavior are sometimes post hoc rationalizations. This means that they are not the results of a transparent access to the inner workings of our mental process, but explanations constructed after the fact to account for our behavior. If so, one can wonder whether participants' answers to Kohlberg's dilemmas really reflect the processes underlying their moral judgment: Rather, they might only reflect how good they are at justifying their answers, which is a totally different cognitive process and ability. To put it otherwise, Kohlberg's method fails to distinguish between the operation of *forming* a moral judgment and the operation of *justifying* a moral judgment. For all we know, both can develop at very different speed.

Finally, another shortcoming of Kohlberg's approach is that it considers the *content* of moral judgments to be irrelevant: The only thing that counts is not people's judgments (what they take to be right or wrong) but the justifications they offer for them. On the one hand, this is actually good scientific methodology: This shows Kohlberg's refusal to consider certain moral judgments as superior to others and thus prevents researchers from imposing their own moral values and ideals as the apex of moral development. We thus avoid the risk of having moral psychology biased by a particular moral worldview. On the other hand, this refusal to take the content of moral judgments into account made Kohlberg's approach hard to connect with the growing body of evolutionary approaches to morality. Indeed, these approaches, because they considered that morality primarily evolved for cooperation, made the prediction that moral development would be marked and structured by the presence of certain themes and norms (reciprocal cooperation rules, tit for tat, etc.). These were predictions that Kohlberg's approach was not well suited to test.

It is for these reasons, among others, that a new paradigm for the psychological study of moral judgment emerged at the beginning of the twenty-first century. It began mainly with Jonathan Haidt's influential 2001 paper, 'The emotional dog and its rational tail: a social intuitionist approach to moral judgment.' Here, Haidt advances the main tenets of intuitionist models of moral judgments:

1. Moral judgments are (mainly) the product of automatic and unconscious cognitive processes to which we have no introspective access. They are *intuitions*.
2. Reasoning actually plays little role in the production of moral judgment. Rather, reasoning comes after, as a means to justify our intuitions to others. Moral reasoning is mainly a post hoc *rationalization*.

These are the lines along which the psychological study of moral judgment developed in the past years. As we will now see, they created the conditions for the contemporary success of the *Trolley Problem*, the fate of which is intimately connected to the rise of intuitionist approaches to moral judgment.

The Trolley Problem as a Flagship for Intuitionist Approaches to Moral Judgment

The intuitionist approach to moral judgment earned rapid success. There are many reasons why, among which its continuity with other findings and paradigms in cognitive and social psychology. However, one reason for this success is that it was also able to provide striking examples of people being unable to justify their own moral judgment. These cases, which Haidt came to call *moral dumbfounding*, were taken to support intuitionist approaches to moral judgment: Indeed, the fact that people can judge something to be wrong (or right) without knowing *why* seems to indicate that they indeed have no access to the psychological processes giving rise to these judgments.

Here is one famous case of moral dumbfounding, put forward by Haidt in his 2001 paper:

JULIE & MARK – Julie and Mark are brother and sister. They are traveling together in France on summer vacation from college. One night they are staying alone in a cabin near the beach. They decide that it would be interesting and fun if they tried making love. At the very least it would be a new experience for each of them. Julie was already taking birth control pills, but Mark uses a condom too, just to be safe. They both enjoy making love, but they decide not to do it again. They keep that night as a special secret, which makes them feel even closer to each other. What do you think about that? Was it OK for them to make love?

And here is how Haidt describes participants' reactions to this case (Haidt 2001, p.814):

Most people who hear the above story immediately say that it was wrong for the siblings to make love, and they then begin searching for reasons. They point out the dangers of inbreeding, only to remember that Julie and Mark used two forms of birth control. They argue that Julie and Mark will be hurt, perhaps emotionally, even though the story makes it clear that no harm befell them. Eventually, many people say something like, "I don't know, I can't explain it, I just know it's wrong." But what model of moral judgment allows a person to know that something is wrong without knowing why?

Because participants still judge that it was wrong for Julie and Mark to have sex together, even when all the reasons they explicitly produce in favor of this judgment have been rebuked, Haidt takes this to show that the *reasons* participants put forward have nothing to do with the actual *cause* of their judgments. To put it otherwise, they ‘know’ it was wrong for Julie and Mark to have sex together, but they don’t know *why*. This is a case of moral dumbfounding.

Though famous and striking, this example suffers from several problems. First and foremost, the corresponding study was never published—its results are only mentioned in Haidt (2001), and the procedure is never fully described. As it turns out, one published attempt at replicating this study found (1) that most participants did not consider that it was wrong for Julie and Mark to have sex together and (2) that the few participants who did were able to produce sound and cogent argument for their judgment (Royzman et al. 2015).

Second, even if we ignore the results of this replication, there are several conceptual problems with what Haidt considers good arguments in favor of condemning Julie and Mark’s sexual relationship. Indeed, Haidt seems to think that the fact that there was no harmful consequence for them means that their act cannot be rationally condemned on the basis of harm-related considerations. However, it is perfectly rational to morally condemn someone for acting in a way that *could* have inflicted harm, even if it did not: If I drive way above the speed limit, putting people at risk, then I deserve blame for being reckless, even if I am lucky enough not to harm anyone. Similarly, it is not because Julie and Mark’s conduct leaves them psychologically unharmed at the end that they cannot be condemned for taking a huge risk. Thus, it is not because participants justify their condemnation with appeal to harm that they are only providing post hoc justifications. Moreover, participants might feel forced to produce such justifications because Haidt seems to consider that saying that Julie and Mark’s action is wrong simply because it is incest does not count as a proper and legitimate justification. But saying that (1) incest is wrong, (2) Julie and Mark’s sexual relationship counts as incest and thus that (3) Julie and Mark’s sexual relationship is wrong seems to be a perfectly sound argument. Thus, it might not be dumbfounding to simply claim that what Julie and Mark did is wrong because it counts as incest.

Fortunately, intuitionist models of moral judgment have a second flagship example that might prove more convincing. This example is known as the *Trolley Problem* and is composed of a pair of cases. Here is the first case (as presented in Hauser et al. 2007):

Denise – Denise is a passenger on a train whose driver has fainted. On the main track ahead are 5 people. The main track has a side track leading off to the left, and Denise can turn the train on to it. There is 1 person on the left hand track. Denise can turn the train, killing the 1; or she can refrain from turning the train, letting the 5 die.

Is it morally permissible for Denise to turn the train?

In this case, Hauser and his colleagues (2007) found that 85% of participants answered ‘YES,’ that it was indeed permissible for Denise to turn the train. Let’s now compare these results with participants’ answers to the second case:

Frank – Frank is on a footbridge over the train tracks. He sees a train approaching the bridge out of control. There are 5 people on the track. Frank knows that the only way to stop the train is to drop a heavy weight into its path. But the only available, sufficiently heavy weight is 1 large man, also watching the train from the footbridge. Frank can shove the 1 man onto the track in the path of the train, killing him; or he can refrain from doing that, letting the 5 die.

Is it morally permissible for Frank to shove the man?

In this case, only 12% of participants answered that it was morally permissible for Frank to shove the man. This is a huge difference from the 85% who found acceptable to turn the train in the first case. However, how are we to explain this difference? It seems that both cases involve causing the death of one person to save five. So, why treat them differently?

What Hauser and his colleagues did was take a subset of participants who answered ‘YES’ to the first case and ‘NO’ to the second and ask them to justify their answers. They then coded these justifications as (1) sufficient justifications, (2) insufficient justifications and (3) discountable objections. Sufficient justifications were justifications that pointed out at least one difference between the two cases (e.g., that there was a physical contact between the agent and his victim in the *Frank* case but not in the *Denise* case). Insufficient justifications were justifications that failed to do so (e.g., justifying one’s answer to the *Denise* case by pointing out that 5 lives are more important than 1, but failing to explain why this no longer matters in the *Frank* case). Finally, discountable objections included both blank answers and answers that made unwarranted assumptions about the scenarios (e.g., that a man’s body cannot stop a train). Once discountable objections were excluded, only 30% of remaining justifications counted as sufficient justifications.

Thus, what these results suggest is that participants can have very strong moral responses to the *Denise* and *Frank* cases but not be able to justify these responses: They cannot explain why they handle the two cases differently. This goes to show that participants’ answers to these cases are driven by features they are not aware of, and the result of psychological processes they do not have introspective access to. This pair of cases, the *Trolley Problem*, is thus a perfect illustration of the intuitionist approaches to moral judgment. It is then no wonder that its appearance and success in the psychological literature coincides with the rise of these approaches.

The Trolley Problem: Origins

Philippa Foot

But the Trolley Problem did not begin its career in the field of psychology. It was first a philosophical puzzle. It made its debut in a 1967 paper by the philosopher Philippa Foot, entitled ‘The problem of abortion and the doctrine of double effect.’

As indicated by its title, the paper deals first and foremost with the moral permissibility of abortion, and mainly with the idea that abortion is morally

permissible when (1) it is performed to save the mother's life and (2) when the (potential) child's death is not a means to save the mother's life, but an unfortunate side effect of the operation required to save the mother's life. This idea is a direct consequence of a moral principle popular in Catholic theology and supposedly put forward by Thomas Aquinas: the *doctrine of double effect*, according to which bad consequences can be brought about to secure a greater good only if they are an unintended side effect of securing this good, and not a means that is therefore intended (see McIntyre 2014 for an introduction).

According to Philippa Foot, this principle is somewhat intellectually attractive. To illustrate its intuitiveness, Foot puts forwards the two following cases:

TROLLEY DRIVER – ‘the driver of a runaway tram (...) 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.’

Here, we recognize the *Denise* case we encountered in the previous section. Foot contrasts this case with the following:

MAGISTRATE & THE MOB – ‘Suppose that a judge or magistrate is faced with rioters demanding that a culprit be found for a certain crime and threatening otherwise to take their own bloody revenge on a particular section of the community. The real culprit being unknown, the judge sees himself as able to prevent the bloodshed only by framing some innocent person and having him executed.’

According to Foot, we have very different intuitions about both cases, and the doctrine of double effect seems to be the best explanation for this difference:

The question is why we should say, without hesitation, that the driver should steer for the less occupied track, while most of us would be appalled at the idea that the innocent man could be framed. It may be suggested that the special feature of the latter case is that it involves the corruption of justice, and this is, of course, very important indeed. But if we remove that special feature, supposing that some private individual is to kill an innocent person and pass him off as the criminal we still find ourselves horrified by the idea. The doctrine of double effect offers us a way out of the difficulty, insisting that it is one thing to steer towards someone foreseeing that you will kill him and another to aim at his death as part of your plan.

Foot then presents another pair of cases that seemingly supports the doctrine of double effect:

DRUG SHORTAGE – ‘We are about to give a patient who needs it to save his life a massive dose of a certain drug in short supply. There arrive, however, five other patients each of whom could be saved by one-fifth of that dose. We say with regret that we cannot spare our whole supply of the drug for a single patient, just as we should say that we could not spare the whole resources of a ward for one dangerously ill individual when ambulances arrive bringing in

victims of a multiple crash. We feel bound to let one man die rather than many if that is our only choice.’

SURGEON – ‘We can suppose, similarly, that several dangerously ill people can be saved only if we kill a certain individual and make a serum from his dead body. (These examples are not over-fanciful considering present controversies about prolonging the life of mortally ill patients whose eyes or kidneys are to be used for others.) Why cannot we argue from the case of the scarce drug to that of the body needed for medical purposes?’

According to Foot, we judge it acceptable to let one person die in the first case, but not to kill one person in the second case. But what is the difference between the two cases? Here again, the doctrine of double effect seems to provide a straightforward explanation: Letting one person die in the first case is just a side effect, while killing a person in the second case constitutes a means. However, Foot ultimately rejects the doctrine of double effect, on the grounds of our intuitions about the following case:

FUMES – ‘Suppose, for instance, that there are five patients in a hospital whose lives could be saved by the manufacture of a certain gas, but that this inevitably releases lethal fumes into the room of another patient whom for some reason we are unable to move. His death, being of no use to us, is clearly a side effect, and not directly intended.’

According to Foot, it is clear that is not morally acceptable to manufacture the gas in this case, even if the death of the patient would be a mere side effect of our action. If so, this case would constitute a counterexample to the doctrine of double effect. But would people actually share Foot’s intuition? To see whether it is the case, I recruited 50 US residents through Amazon Mechanical Turk (paid \$0.30 each, 30% women, $M_{\text{age}} = 37.8$, $SD_{\text{age}} = 11.3$) and gave them the following scenario:

Suppose that there are five patients in a hospital whose lives could be saved by the manufacture of a certain gas, but that this inevitably releases lethal fumes into the room of another patient whom for some reason doctors are unable to move.

Question. Is it morally acceptable for doctors to manufacture this gas, thereby saving five patients but killing one? (YES/NO)

Overall, 48% of participants answered ‘YES’ and 52% answered ‘NO.’ As we can see, though participants’ answers are not as straightforward as in *Frank’s* case (in which saving five persons required shoving one person into the track), people seem indeed reluctant to endorse manufacturing the gas. As suggested by Foot, this case seems to constitute a counterexample to the doctrine of double effect.

But, if the doctrine of double effect cannot accommodate our intuitions about such cases, what is the explanation? Foot puts forward the traditional distinction between ‘negative duties’ that forbid us to perform certain actions (e.g., you shall not steal) and ‘positive duties’ that require us to perform certain actions (e.g., you must help those in need). According to Foot, the principle according to which we

must choose the course of action that produces the greater good only holds when comparing duties of the same nature. But when the choice is between respecting a negative duty and a positive one, the negative duty always has the priority. For example, in *DRUG SHORTAGE*, we have the choice between two positive duties (helping five patients or helping one patient). In this case, it is thus morally acceptable to seek the greater good and help five patients, even if it means letting one die. But, in *SURGEON*, our positive duty (helping five patients) conflicts with a negative duty (not killing one patient). This is why it becomes unacceptable to seek the greater good. The same is true for *MAGISTRATE & THE MOB* and *FUMES*.

But what about *TROLLEY DRIVER*? In this case, according to Foot, we have the choice between two negative duties. Indeed, the trolley driver can either kill one person or kill five persons. Because the conflict is between two duties of the same nature, it is acceptable to seek the greater good and thus acceptable to divert the trolley.

Judith Jarvis Thomson

With the *TROLLEY DRIVER* case, Foot gave us the first half of the Trolley Problem. But the second half is due to another famous moral philosopher, Judith Jarvis Thomson. In response to Foot, Thomson developed a new series of cases, in two different papers (Thomson 1976, 1985).

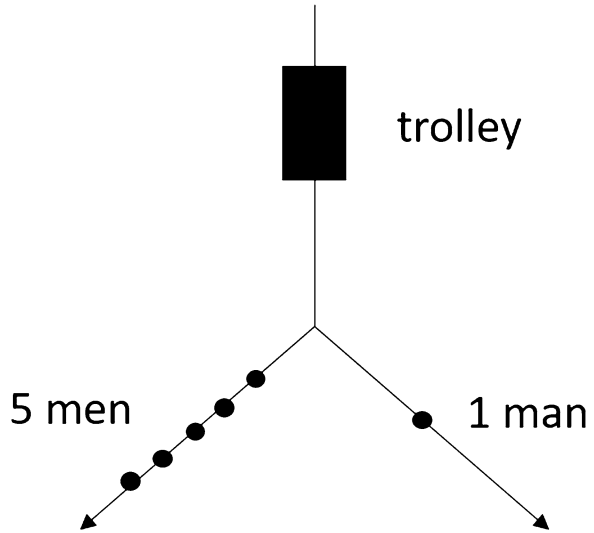
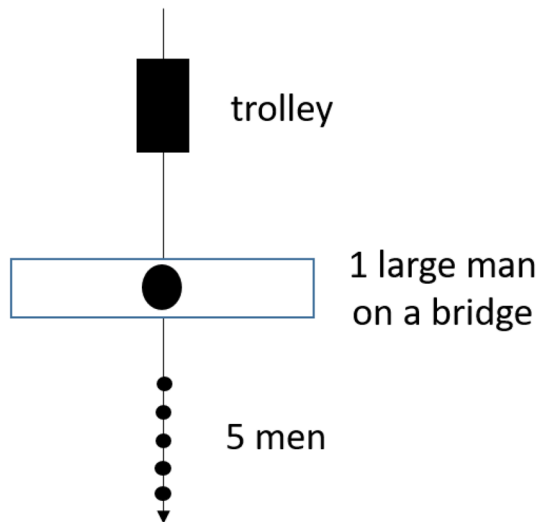
First, Thomson rejects Foot's idea that our intuitions about the *TROLLEY DRIVER* can be explained by Foot's solution. To prove it, she slightly modifies the case to create the following variation (Fig. 1):

BYSTANDER – 'In that case you have been strolling by the trolley track, and you can see the situation at a glance: The driver saw the five on the track ahead, he stamped on the brakes, the brakes failed, so he fainted. What to do? Well, here is the switch, which you can throw, thereby turning the trolley yourself. Of course you will kill one if you do. But I should think you may turn it all the same.' (Thomson 1976, p.1397)

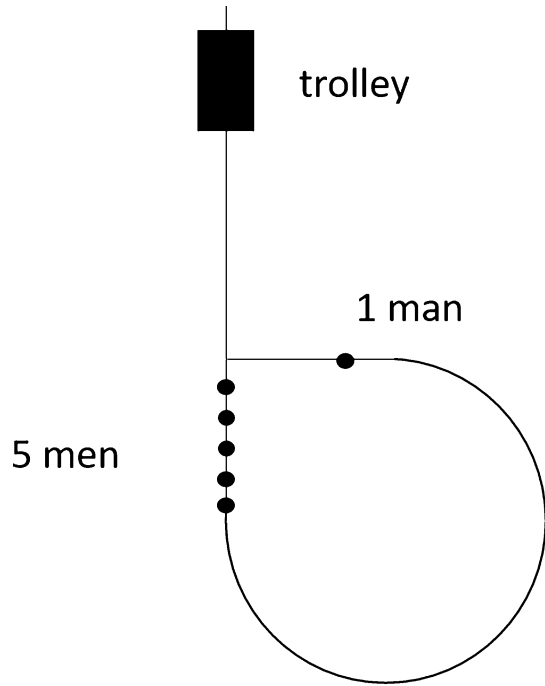
The *BYSTANDER* is the most widespread variation of the side effect part of the Trolley Problem. It has been repeatedly shown that people's intuitions are in line with Thomson's and that most people consider it morally acceptable to divert the trolley in this case. However, because the agent is no longer the trolley's driver, but an innocent bystander, this version of the case pits a negative duty (do not cause the death of one person) against a positive duty (save five persons). Still, contrary to Foot's predictions, we judge perfectly acceptable to cause the death of one person to save five.

It is also Thomson who introduces *FRANK*'s case, better known as the *FOOTBRIDGE* case (Fig. 2), as a perfect contrast to the *BYSTANDER* case and as the most striking way of asking the question: Why is it morally worse to save five people by throwing someone under a trolley than by diverting a trolley into someone? *This* is the Trolley Problem.

However, Thomson agrees with Foot to conclude that the doctrine of double effect is *not* the solution to the Trolley Problem. As a demonstration, she introduces

Fig. 1 BYSTANDER case**Fig. 2** FOOTBRIDGE case

a third case, the LOOP case (Fig. 3), in which the main track (with five men) and the side track (with one man) are in fact the same track that loops on itself. Letting the trolley go on its actual course will lead it to collide with five men, which will stop it and save the one man. However, diverting the trolley on the one man will lead it to collide with the one man, which will stop it and save the five men. According to Thomson, it is acceptable to divert the trolley in this case. However, using the one man to stop the trolley is using him as a means, and not only causing his death as a side effect. If Thomson is right, then this case constitutes another example to the doctrine of double effect.

Fig. 3 LOOP case

But is she right? In order to know, I recruited 50 US residents through Amazon Mechanical Turk (paid \$0.30 each, 46% women, 4% unidentified, $M_{\text{age}} = 35.1$, $SD_{\text{age}} = 10.2$) and gave them the following scenario (together with Fig. 3):

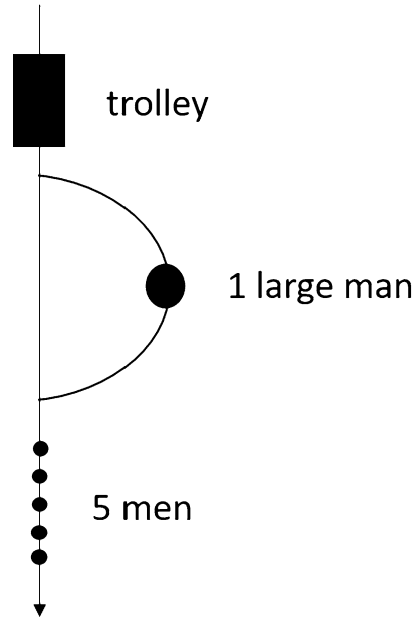
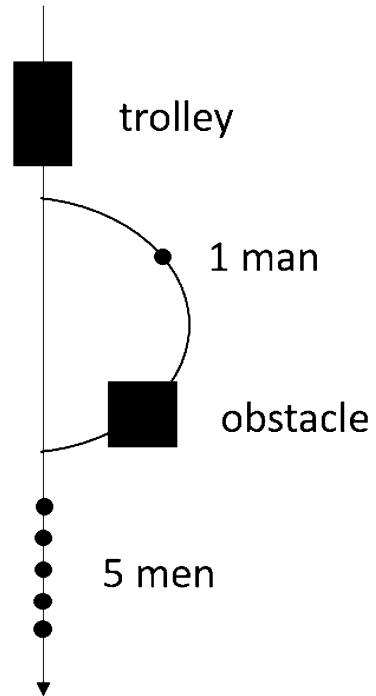
A runaway trolley is heading down a track that loops on itself. If it proceeds on its present course, it will collide with a group of five railway workmen, who will be killed. This will stop the trolley and save a sixth workman who is working further down the track.

Denise is a bystander near the track. If she wants, she can flip a switch that will divert the trolley on its left. If Denise does so, the trolley will collide with the lone workman. This will kill the lone workman, but this will also stop the trolley and save the five workmen.

Question: Is it morally permissible for Denise to divert the trolley? (YES/NO)

In this case, in line with Thomson's intuition, 74% of participants answered 'YES.' Thus, participants seem perfectly willing to killing a person as a means, which goes against the doctrine of double effect.

So, what is Thomson's solution? For Thomson, it is a question of *rights*: Causing the death of one person to save many is morally wrong when doing so violates this person's rights. But when we divert an already existing threat (rather than creating a new threat) onto someone, we do not violate this person's rights. Why? Thomson admits she cannot really tell why, but insists that her proposal is somewhat intuitive.

Fig. 4 NED case**Fig. 5** OSCAR case

From Philosophy to Psychology

In its original formulation, the Trolley Problem is a *normative* question. Both Foot and Thomson assume that it is indeed morally permissible to divert the trolley and morally wrong to push the large man under the trolley. The question is thus: Why? What is the morally relevant difference between the two cases? What are the correct moral rules that account for this normative difference between the two cases?

However, not everyone agrees to take our intuitions on these cases at face value. Indeed, it is possible to argue that our intuitions about such cases are biased and thus do not properly reflect mind-independent moral truths. This is what Peter Unger argues in his book *Living high and letting die: Our illusion of innocence* (Unger 1996). Drawing on several variations of the aforementioned scenarios, Unger argues that the factors that shape our intuitions about these cases are morally irrelevant and thus that these intuitions should not be used as a basis for ethical reflection.

Of course, Unger's thesis is a psychological thesis on the psychological determinants of our answers to such cases. This goes to show that philosophical discussions about the Trolley Problem already integrated a *psychological* approach to the problem, which investigated the causes of our intuitions.¹ This is why, when the time came, it was easy for the problem to make the jump to psychology.

The psychological uses of the Trolley Problem

However, the success of the Trolley Problem in the psychological literature cannot be explained simply by its success as an illustration of the intuitionist approaches to moral judgment. In fact, it has been put to very different uses by psychologists and neuroscientists.

The Trolley Problem as a 'Fruit Fly of the Moral Mind'

We have already seen that intuitionist approach to moral judgment reject introspection and the use of participants' justifications. But, under these conditions, how are we to investigate the unconscious psychological process that underlies our moral judgments? One way is the use of 'minimal pairs,' that is, the use of pairs of scenarios that differ only along a particular dimension. If participants' judgments differ between the two scenarios, then it is evidence that this dimension is taken into account by the psychological process that drives our moral judgments. This is why such scenarios are sometimes called 'probes': They allow us to probe the inner workings of our moral psychology.

This method was already standard in social psychology, for example in the study of the action/omission distinction (see Spranca et al. 1991 for an example). But the Trolley Problem provides a striking example of such a pair: It provides a model for

¹ Actually, in a footnote of his book, Unger tells us that he sought the help of his colleagues from the Psychology Department to put his hypothesis to the test. Unfortunately, he didn't have much success. The times were not ripe for the empirical study of moral intuitions.

future research on the unconscious process guiding moral judgment. Following this model, studies have built similar minimal pairs to investigate the role of factors such as the presence of a contact between the agent and the victim, or the distinction between intended harm and mere side effects (Cushman et al. 2006). Because such scenarios are easy enough to manipulate, and provide us a window into the moral mind, Joshua Greene has described them as ‘fruit flies of the moral mind’ (Greene 2009; see also Cushman and Greene 2012).

For example, it has been argued that part of the Trolley Problem, namely the difference between the *BYSTANDER* and *FOOTBRIDGE* cases, can be explained by the fact that, in line with the doctrine of double effect, people make an implicit difference between harm that is intended as a means and harm that is brought only as a side effect. To test for this and exclude potential confounding factors, Hauser and his colleagues (2007) designed an even more minimal pair than the original one (Figs. 4 and 5):

NED – Ned is walking near the train tracks when he notices a train approaching out of control. Up ahead on the track are 5 people. Ned is standing next to a switch, which he can throw to turn the train on to a side track. There is a heavy object on the side track. If the train hits the object, the object will slow the train down, giving the men time to escape. The heavy object is 1 man, standing on the side track. Ned can throw the switch, preventing the train from killing the 5 people, but killing the 1 man, or he can refrain from doing this, letting the 5 die.

OSCAR – Oscar is walking near the train tracks when he notices a train approaching out of control. Up ahead on the tracks are 5 people. Oscar is standing next to a switch, which he can throw to turn the train on a side track. There is a heavy object on the side track. If the train hits the object, the object will slow the train down, giving the 5 people time to escape. There is 1 man standing on the side track in front of the heavy object. Oscar can throw the switch, preventing the train from killing the 5 people, but killing the 1 man. Or he can refrain from doing this, letting the five die.

As in the original Trolley Problem, one case involves using the victim as a means (in the *NED* case, to slow down the train), and the other involves killing him as a side effect (in the *OSCAR* case, because he unfortunately stands in front of the object used to slow down the train). However, other potential differences have been erased: Neither case involves a contact between the agent and the victim, and the basic action performed by the agent is the same in both cases (throwing the switch). In these conditions, Hauser and his colleagues found that 56% of participants judged that it was morally permissible for Ned to throw the switch, while 72% of participants judged that it was morally permissible for Oscar to throw the switch. This is a much smaller difference than the original difference between *DENISE* and *FRANK*. Moreover, a recent meta-analysis suggests that, even if the means/side effect distinction indeed has an effect on moral judgment, this effect is quite small (Feltz and May 2017). Together, these results suggest that, even if participants are

somewhat sensitive to the difference between means and side effects, this difference cannot explain the whole Trolley Problem.

The Trolley Problem as Proof of a Universal Moral Faculty

Another feature of the Trolley Problem that attracted the attention of psychologists is its apparent universality. Indeed, when it was first tested, the difference between DENISE and FRANK and between the BYSTANDER and FOOTBRIDGE cases was observed in several countries and seemed rather insensitive to demographic factors such as education and religiosity (see Hauser et al. 2007). Moreover, it had also been observed in young children (Pellizzoni et al. 2010).

This is why some have taken the Trolley Problem to be evidence that humans are endowed with an innate, biologically rooted moral faculty. The most famous version of this thesis is probably the ‘Universal Moral Grammar,’ a thesis according to which this moral faculty is structured similarly to our linguistic capacity (Hauser 2006; Mikhail 2011). Though the ‘Universal Moral Grammar’ drew a lot of attention (including criticisms) for a time (see for example Mikhail 2007; Dupoux and Jacob 2007), it has mostly disappeared from the psychological scene for now. One reason might be that recent studies have stressed the cultural variability of intuitions about the Trolley Problem (Michelin et al. 2010; Ahlenius and Tännsjö 2012).

Trolleys and the Dual-Process Theory of Moral Judgment

Another domain in which the Trolley Problem has been heavily put to use is the defense of the dual-process theory of moral judgment. The dual-process theory of moral judgment, put forward by philosopher and neuroscientist Joshua Greene, accepts most of the conclusion of the intuitionist approach to moral judgment while giving a larger place to conscious reasoning in our moral lives. According to this model, inspired by the distinction between System 1 (intuitive, unconscious) and System 2 (reflexive, conscious), both intuitions and reasoning can play a role in the formation of moral judgment.

One way to show that this is the case is simply to find a case in which the two systems conflict. According to Joshua Greene, trolley cases are such cases: They pit an emotional, deontological response (i.e., do not kill) against a reflexive, utilitarian response (i.e., it is acceptable to kill one person to save many). Our decisions in trolley cases would thus reveal the structure of moral cognition, by revealing the conflict between the two systems.

Joshua Greene was one the first researcher to use trolley cases for a psychological purpose. In 2001, together with his colleagues, he published in *Science* a paper using fMRI to investigate the neural basis of moral judgment about moral dilemmas (i.e., cases involving harming someone to bring about a greater good). Using standard trolley cases (such as BYSTANDER and FOOTBRIDGE), they were able to draw a distinction between ‘personal’ scenarios (in which the harm is brought directly to the victim, as in FOOTBRIDGE) and ‘impersonal’ scenarios (in which the harm is brought indirectly to the victim, as in BYSTANDER). They found that personal

scenarios (that elicit more deontological responses) elicited more activation in emotion-related brain areas (medial frontal gyri, angular gyri) than impersonal scenarios (that elicit more utilitarian responses), and that impersonal scenarios elicited more activation in areas involved in higher cognition, reasoning and cognitive control (right middle frontal gyrus, left and right parietal lobes). This suggests that moral dilemmas involve a conflict between higher-order cognition (leaning toward a utilitarian calculation) and emotional responses (triggered by the idea of harming someone).

Though influential, this study has serious shortcomings. Because ‘personal’ and ‘impersonal’ dilemmas are not matched in minimal pairs (with the exception of the BYSTANDER/FOOTBRIDGE pair), one can doubt that the differences observed in brain activation are due to the contrast between ‘direct’ and ‘indirect’ harm. Indeed, in this study, most personal dilemmas involve violent death or outrageous deeds (murder, rape), while impersonal dilemmas tend to involve less violent outcome (cheating on one’s taxes, stealing a wallet). It is thus not very surprising that personal scenarios tend to elicit stronger emotional reactions. Similarly, impersonal scenarios tend to include more dilemmas making use of probability and percentages than personal scenarios. It is thus not very surprising that impersonal scenarios tend to involve higher-order cognition more than personal ones.

However, beyond brain imagery, there is plenty of evidence for the dual-process theory of moral judgment (see Greene 2014 for a review). These can be classified in three categories. First, correlational data show certain relationships between certain individual traits and the tendency to give utilitarian responses. For example, it has been shown that individuals who score high on psychopathy scale tend to give more utilitarian responses to these dilemmas, which suggests that an absence of emotional response does favor utilitarian judgment (Bartels and Pizarro 2011). Second, experimental studies show that undermining reasoning and higher-order cognition tends to favor deontological responses. For example, one study found that forcing participants to answer quickly increased the rate of deontological responses, while forcing them to take the time to think increased the rate of utilitarian responses (Suter and Hertwig 2011). Third, neuropsychological evidence suggests that patients suffering from emotional impairment following damage to certain brain areas are more likely to give utilitarian responses (Koenigs et al. 2007). Whether these are convincing evidence is up for debate, but there is no doubt that the Trolley Problem plays a major role in all these studies.

The Trolley Problem as a Standard Stimulus in Moral Psychology

Finally, it is important to note that the Trolley Problem has gained such a notoriety from these different uses that he has become a standard stimulus in moral psychology. To put it otherwise, responses to the Trolley Problem are used as standard examples of moral judgment in studies that could very well use completely different stimuli. For example, it has been used in studies about the effect of sleep deprivation on moral judgment (Killgore et al. 2007) or the effect of alcohol on moral judgment (Duke and Bègue 2015). It has also been widely used in studies about order effects and framing effects in moral judgment (e.g., Nadelhoffer and

Feltz 2008). In all these instances, trolley cases are used in studies that do not aim at understanding the nature of the Trolley Problem (and discovering the reasons why we treat BYSTANDER and FOOTBRIDGE differently) and it is not clear that the same research could not have been conducted by using radically different stimuli.

That trolley cases have become a widespread stimulus in the study of moral judgment can be seen as worrying, given that we still have no clear idea of what is going on in such cases and that the cases have a certain air of artificiality. However, it is interesting to note that the first occurrence of trolley cases in the psychological literature belongs to this category: In 1993, Petrinovich and his colleagues published a paper using trolley cases to investigate different factors shaping moral judgment, like one's genetic closeness to the victim. In 1996, Petrinovich and O'Neill also used trolley cases to investigate wording and framing effects on moral judgments. However, in both cases, it is clear that trolley cases were used as a means to investigate something else and were not the focus of the research.

Has the Trolley Problem Been Solved?

Nowadays, trolley cases are mostly used in two ways by psychologists: either as a way to test the dual-process theory of morality or as a standard stimulus used to investigate something else. Very few papers actually aim to uncover the factors that lead us to treat BYSTANDER and FOOTBRIDGE differently. Does it mean that the Trolley Problem has already been solved? Let's review the different hypotheses that have been put forward.

Doctrine of Double Effect

As mentioned earlier, both Hauser (2006) and Mikhail (2011) argued that the difference between the two cases could be explained by the doctrine of double effect: FOOTBRIDGE involves treating the victim as a means, while the victim's death in BYSTANDER is a mere side effect. However, there are reasons to doubt that this explanation is sufficient. As we saw earlier in the contrast between NED and OSCAR, the distinction between means and side effects seems to explain very little. Moreover, other studies found no significant differences between the two cases, with most participants judging the action to throw the switch morally acceptable in both cases (Greene et al. 2009). It thus seems that, as Foot and Thomson already argued years ago, the doctrine of double effect cannot explain our intuitions about trolley cases.

Doctrine of Personal Force

Let's then turn to another explanation: the 'doctrine of personal force.' The expression has been coined by Greene and his colleagues (2009) as a label for the idea that the means/side effect distinction actually matters, but only when it interacts with another factor: personal force. More precisely, their claim is the following: People judge it morally unacceptable to kill one person to save many

only when (1) the death of the victim counts as a means (rather than a side effect) and (2) the agent kills the victim by acting directly on him or her, that is, by applying a ‘personal force’ on him or her.

To test this claim, Greene and his colleagues have designed an experiment in which they vary independently two factors: the application of personal force and the means/side effect distinction. The NED and OSCAR cases represent the two cases in which no personal force is applied, one being an instance of killing as a means and the other an instance of killing as a side effect. The case involving both personal force and killing as a means is the OBSTACLE PUSH case, in which the agent, located on a bridge, must throw a switch to divert the train and save five people on the tracks (as in BYSTANDER) but is separated from the switch by a man. To reach the switch, the agent must throw the man out of the way and thus off the bridge. Finally, in the case involving personal force and killing as a side effect (OBSTACLE COLLIDE), the situation is similar, except that the agent can simply run toward the switch, but as a side effect will collide with the man, knocking him off the bridge. What Greene and his colleagues found was that participants tended to judge it morally acceptable to kill one person in the NED, OSCAR and OBSTACLE COLLIDE cases, while they tended to judge that unacceptable in the OBSTACLE PUSH case. Thus, killing was judged unacceptable only when there were both personal forces and the use of killing as a means.

However, and despite these striking results, it is not clear that the doctrine of personal force is an adequate solution to the Trolley Problem, as it faces obvious counterexamples. One of these counterexamples is Foot’s FUMES case, in which there is no personal force applied. Another counterexample would be the following: Imagine a case similar to the FOOTBRIDGE case, but in which the agent is not on the footbridge, near the large man. Rather, his only way of pushing the fat man onto the tracks is to shoot him from a distance with a rifle. This is a case in which no ‘personal force’ is applied (pulling the trigger is no more personal than throwing a switch), but it seems no different (and not more acceptable) than the original FOOTBRIDGE case. We thus need another solution.

Agent Versus Patient Intervention

One last solution that has been proposed to the Trolley Problem relies on a distinction between two loci of interventions: intervening on the threat (‘agent intervention’) or intervening on the victim itself (‘patient intervention’). According to Waldmann and Dieterich (2007), the psychological processes underlying our moral responses to trolley cases suffer from ‘intervention myopia’: They only react negatively when the agent acts directly on the victim, and stay silent when the agent acts on the threat to redirect it. This is why we condemn the agent in the FOOTBRIDGE but not in the BYSTANDER case.

Waldmann and Dieterich illustrate this principle throughout several studies. For example, they show that people judge it unacceptable to push a bus on the tracks to stop a runaway train threatening to kill people, but judge acceptable to divert the same train on side track where it will collide with a bus already present on the tracks. Their account also avoids some of the difficulties met by the doctrine of

double effect and doctrine of personal force. However, it also has its problem. For example, it cannot account for Foot's FUMES. Also, it cannot account for the following case:

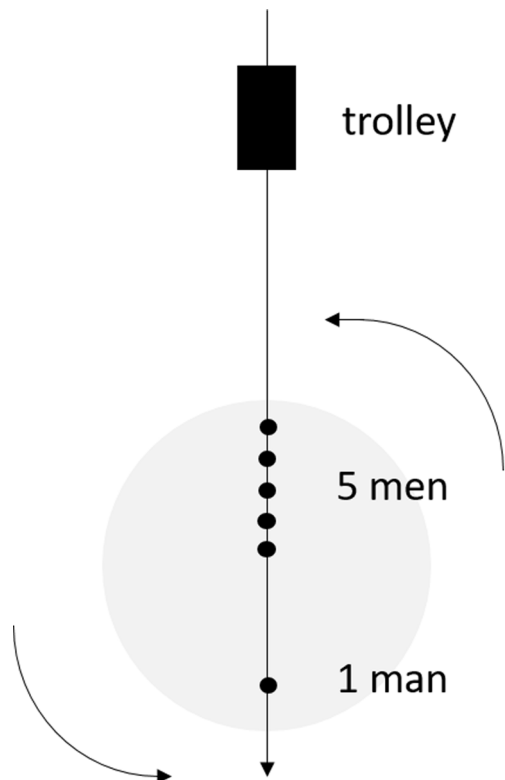
MOVING TRACKS – A runaway trolley is heading down a track. If it proceeds on its present course, it will collide with a group of five railway workmen, who will be killed. This will stop the trolley and save a sixth workman who is working further down the track.

Denise is a bystander near the track. If she wants, she can flip a switch that will quickly rotate the track on which the six workmen are situated, placing the lone workman between the trolley and the five other workmen. If Denise does so, the trolley will collide with the lone workman. This will kill the lone workman, but this will also stop the trolley and save the five workmen.

Question: Is it morally permissible for Denise to rotate the track? (YES/NO)

I presented 50 US residents recruited through Amazon Mechanical Turk (paid \$0.30 each, 40% women, 2% unidentified, $M_{\text{age}} = 33.0$, $SD_{\text{age}} = 10.3$) with this case (together with Fig. 6). 64% of participants judged that it was morally permissible for Denise to rotate the track. However, by rotating the track, Denise is precisely acting on the victim himself (changing his position) rather than on the threat. Thus, these results contradict Waldmann and Dieterich's explanation.

Fig. 6 MOVING TRACKS case



A New Hypothesis: Close Versus Distant possibilities

The Trolley Problem is thus far from being solved: We still don't know what explains the difference between the *BYSTANDER* and the *FOOTBRIDGE* cases. Nevertheless, and despite the Trolley Problem's fame and widespread utilization in psychological research, very few researchers are currently investigating the factors that might explain this difference. This is unfortunate, as understanding the determinants of this difference could teach us a lot about the mechanisms of human moral cognition. This is why I hope the present article will rekindle the motivation to investigate and research the Trolley Problem.

However, in the hope of fostering further research, before ending I would like to set out my own hypothesis. I have no evidence yet for this hypothesis, and it might turn as wrong as the others, but false hypotheses may sometimes have the benefit of fostering discussion and bringing new important questions to light.

So, what is my hypothesis? It relies on the idea that we consider some possibilities to be closer to the actual world than others. That it rains tomorrow is a close possibility: This seems likely. That it rains for 40 days and nights without stopping is less likely: It is a remote possibility. And that it rains (literally) cats and dogs seems almost impossible: It is a very remote, inaccessible possibility.

My proposal is thus the following: That killing one person to save many is more unacceptable when it was unlikely (i.e., a remote possibility) that the victim would die this way. In the *BYSTANDER* case, the victim ends up dying in a train accident. Given that the victim was working on the tracks, this kind of death was a live possibility. The same is true in the *LOOP*, *NED* and *OSCAR* cases. This is why we judge it somewhat acceptable to have the victim being run over by a train in those cases. However, in the *FOOTBRIDGE* case, the person is standing on a bridge, outside the tracks. It is thus harder to come up with a sequence of events that would result in him being crushed by a train. And it is thus unacceptable to kill the person. The same is true in the case of *FUMES*: We tend to think it unacceptable to bring about the patient's death, because dying from poisonous gas is not something we typically expect to happen in a hospital.

This hypothesis can explain the apparent relevance of factors such as killing as means, the use of personal force or acting on the victim (rather than on the threat): Cases in which the death of the victim can be obtained simply as a side effect without having to use personal force on the victim himself will be more likely to be cases in which the victim is already in a precarious situation, and thus closer to possible worlds in which he dies in the intended way. For example, a variant of the *FOOTBRIDGE* in which there is no use of personal force is the following: The agent is not on the footbridge, near the large man, but can activate at distance a trapdoor situated just below the large man, that will drop him onto the tracks. In this case, people tend to think it is morally acceptable to activate the trap door to have the large man fall on the tracks (Greene et al. 2009). This goes against both the doctrine of double effect and the idea that intervening on the victim himself leads to moral condemnation. But this is easily explained by my account: Having the large man standing on a trapdoor immediately puts him in a situation where his falling on the tracks becomes a salient, close possibility.

But how could we put this hypothesis to test? Here is one possibility. Another variant of the *FOOTBRIDGE* has the bridge already planted with explosives. In this

case, the agent can detonate the explosives to destroy the bridge and have the rubble stop the train—it will only kill the man standing on the bridge as a side effect. In this case, most people consider it acceptable to destroy the bridge (Mikhail 2007), and this is in line with my account: A person standing on a bridge full of explosives is likely to die in the bridge’s explosion. However, what my hypothesis predicts is that it won’t be morally acceptable to blow up the bridge if the bridge was not already likely to explode in the absence of the agent’s intervention. We could thus imagine a variation of the case in which the agent does not explode the bridge by using explosives that are already present, but by using a bazooka or some kind of missile he has in his possession. In this case, my account predicts that it would be unacceptable to blow up the bridge and kill the person standing on it, even if his death would only be side effect brought without use of personal force.

Conclusion

In this paper, my aim was to give an overview of the Trolley Problem, of its philosophical origins, and of its different uses in the field of moral psychology. It was also to argue that, despite its popularity and widespread utilization in psychological research, few researchers actually tried to directly solve it and that we are still ignorant of the real factors guiding our responses to trolley cases. Against the idea that Trolley Problems were just a fad, a speculative bubble doomed to burst, I hope by the present paper to rekindle researchers’ interest in investigating the sources of our intuitions in trolley cases.

Supplementary Materials

Data for the three studies presented in this paper as well as publicly available versions of the figure can be found on the *Open Science Framework* at the following address: <https://osf.io/zy4st/>.

References

- Ahlenius, H., & Tännsjö, T. (2012). Chinese and Westerners respond differently to the trolley dilemmas. *Journal of Cognition and Culture*, 12(3–4), 195–201.
- Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121(1), 154–161.
- Cushman, F., & Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social Neuroscience*, 7(3), 269–279.
- 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.
- Duke, A. A., & Bègue, L. (2015). The drunk utilitarian: Blood alcohol concentration predicts utilitarian responses in moral dilemmas. *Cognition*, 134, 121–127.
- Dupoux, E., & Jacob, P. (2007). Universal moral grammar: A critical appraisal. *Trends in Cognitive Sciences*, 11(9), 373–378.

- Feltz, A., & May, J. (2017). The means/side-effect distinction in moral cognition: A meta-analysis. *Cognition*, 166, 314–327.
- Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, 5, 5–15.
- Greene, J. (2009). Fruit flies of the moral mind. In M. Brockman (Ed.), *What's next? Dispatches on the future of science* (pp. 104–115). Toronto: Random House of Canada.
- Greene, J. D. (2014). Beyond point-and-shoot morality: Why cognitive (neuro) science matters for ethics. *Ethics*, 124(4), 695–726.
- Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111(3), 364–371.
- 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(5537), 2105–2108.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.
- Hauser, M. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York: Ecco/HarperCollins Publishers.
- Hauser, M., Cushman, F., Young, L., KangXing Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind and Language*, 22(1), 1–21.
- Killgore, W. D., Killgore, D. B., Day, L. M., Li, C., Kamimori, G. H., & Balkin, T. J. (2007). The effects of 53 hours of sleep deprivation on moral judgment. *Sleep*, 30(3), 345–352.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, 446(7138), 908–911.
- Kohlberg, L. (1963). The development of children's orientations toward a moral order. *Human Development*, 6(1–2), 11–33.
- Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. *Theory into Practice*, 16(2), 53–59.
- McIntyre, A. (2014). Doctrine of double effect. In: E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy* (Winter 2014). Stanford: Metaphysics Research Lab, Stanford University.
- Michelin, C., Pellizzoni, S., Tallandini, M. A., & Siegal, M. (2010). Should more be saved? Diversity in utilitarian moral judgment. *Journal of Cognition and Culture*, 10(1), 153–169.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences*, 11(4), 143–152.
- Mikhail, J. (2011). *Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment*. Cambridge: Cambridge University Press.
- Nadelhoffer, T., & Feltz, A. (2008). The actor–observer bias and moral intuitions: Adding fuel to Sinnott-Armstrong's fire. *Neuroethics*, 1(2), 133–144.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.
- Pellizzoni, S., Siegal, M., & Surian, L. (2010). The contact principle and utilitarian moral judgments in young children. *Developmental Science*, 13(2), 265–270.
- Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology*, 17(3), 145–171.
- Petrinovich, L., O'Neill, P., & Jorgensen, M. (1993). An empirical study of moral intuitions: Toward an evolutionary ethics. *Journal of Personality and Social Psychology*, 64(3), 467.
- Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, 10(4), 296.
- Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, 27(1), 76–105.
- Suter, R. S., & Hertwig, R. (2011). Time and moral judgment. *Cognition*, 119(3), 454–458.
- Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist*, 59(2), 204–217.
- Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal*, 94(6), 1395–1415.
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge: Cambridge University Press.
- Unger, P. K. (1996). *Living high and letting die: Our illusion of innocence*. USA: Oxford University Press.
- Waldmann, M. R., & Dieterich, J. H. (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.