

The Differences and Similarities of Different Philosophers on Tackling the Trolley Dilemma

FengYu Yang

*Germantown Friends School, Philadelphia, 19144, United States.
Email: johnyang515410@gmail.com*

ABSTRACT

The Trolley Dilemma, since originally brought up by Philippa Foot in 1967, has caused a series of debates among philosophers and scholars from various schools of thought. Though the academic field has yet to step foot on a common ground, The Trolley Dilemma has consistently induced many discussions when taking context of factors in the developing world into consideration, resulting in numerous variations of the original dilemma, such as Transplant, Bystander, and Fatman. Despite their distinct forms, they all share the same struggle for the decision makers – the trolley driver, the bystander, or other roles: whether to let five people die or sacrifice one equally innocent person to save the five. In this paper, our philosophers’ (Judith Jarvis Thomson, Joshua D. Greene, Peter Singer, Frances M. Kamm) view on the Trolley Dilemma is being critically analyzed, compared and contrasted, and ends upon the suggestion of potentially “walking away” as a form of nonaction which eradicates the ought. Through comparing the similarities and differences, strengths and drawbacks among theories proposed by great minds, I plan to state that there will never be an unanimously agreed upon solution to this dilemma. I regard the papers being in a “debate” because there will only be variants after variants created by future philosophers in order to refute the past essays, leading to nowhere apart from suffocating humanity and themselves within it. Nonetheless, although the dilemma may never be solved, from the unstoppable debates, implied ideas may be exerted, leading to advancements on the field of morality. Comparatively, I will contribute to academia by interpreting significant thoughts thoroughly, and through comparing and contrasting from a neutral standpoint, I will bring in new perspectives to approach them. I believe that my paper will give rise to more critical thoughts on this never-ending and controversial philosophical subject and invite more who are interested to participate in this debate.

Keywords: *Trolley Dilemma, The Trolley Problem, Judith Jarvis Thomson, Joshua D. Greene, Peter Singer, Frances M. Kamm*

1. INTRODUCTION

The Trolley Dilemma should be the most well-known set of philosophical dilemmas, (the original) with simply one trolley, one decision, two tracks, and six innocent people, it has certainly been the signifying case of moral philosophy since it was brought up in Philippa Foot’s paper (*The Problem of Abortion and the Doctrine of the Double Effect*) from 1967. Over the past years, numerous philosophers have been attempting to solve the dilemma with their supposedly personal answers for it. Yet, with no exceptions, the solution (usually a theory) would be refuted by the introduction of another variation of the Trolley.

Even though The Trolley Dilemma was addressed by a series of philosophers(scholars) in the past, do those

series share any similarities, differences, or have a communal aspect which they have not touched upon? Hence, in this paper, I will start by briefly examining several primary or secondary resources related to The Trolley Dilemma and then cross-checking for similarities and differences, then end with an aspect of “walking away” which I believe no philosopher has yet touched upon.

Later in the text, four philosophers would be consecutively introduced due to their work on The Trolley Dilemma: Judith Jarvis Thomson, Joshua D. Greene, Peter Singer, and Frances M. Kamm. These four philosophers’ works were chosen for the great variety they incorporate while looking upon the Trolley Dilemma, including the notion of rights, neurology, history, and the notion of intention.

Before introducing any paper, I would like to start by talking about the trolley dilemma and its variants and to differentiate the trolley dilemma from The Trolley Problem.

Trolley, the original case that was raised by Foot, and the following variants which were later invented, are generally known as trolley dilemma(s), mentioned, discussed, and studied by numerous scholars around the world. Yet, The Trolley Problem is an elaborated version of the trolley dilemma(s) brought up by Judith Jarvis Thomson, focusing on the issue of why **Bystander** and **Transplant**, two nearly identical variants with similar situations and actions, would lead to drastically different intuitive outcomes [4].

2. JUDITH JARVIS THOMSON AND THE TROLLEY PROBLEM

To begin with, Thomson initiated with the Kantian idea of one should not treat people “as a means only” to save another’s life [4]. This theory seemed to fit with **Transplant** as the surgeon clearly uses the young man as a means to save the others. Yet, Thomson comes up with the variant that I will call **Loop**. Within this variant, the larger workman tied on the loop to stop the trolley would be similarly used as a means for the other workmen, but many would still consider turning the trolley onto the larger workman as permissible, proving the inapplicability of the Kantian theory in the Trolley Problem.

Later, Thomson shifts gears into an emphasis on rights for her solution to The Trolley Problem. She begins with a quote from Dworkin, “Rights ‘trumps’ utilities”, meaning “if one would infringe a right in or by acting, then it is not sufficient justification for acting that one would thereby maximize utility” [4]. Hence, any intentional acts of killing (being an infringement of one’s right to life) would not be permissible, even if the net survival rate would be maximized from the person’s death. With Dworkin’s quote as a basis, Thomson develops her most essential argument of the essay, “distributive exemption”, which permits an already existing threat to be distributed onto the fewer [4]. In **Bystander**, it would be permissible to redirect the trolley because (i) the agent is not infringing anyone’s rights (the agent is dealing with an already existed threat) and (ii) the agent is also distributing the already existing threat from the more to the fewer (five to one). Thomson goes deeper into the idea of infringing rights with **Fatman**. She believes, compared to the agent in **Bystander** not infringing any rights of anybody no matter his action, the rights of the fat man would be infringed if he was pushed off the bridge with no necessity of considering how the fat man would end after the push. Since the agent in **Bystander** may only turn the trolley into different tracks, shoving the fat man off the bridge or even shaking the

handrail (as an attempt to let the fat man fall off the bridge) would not be permissible because the agent would have the means to let the fat man topple of the bridge and thereby infringing his rights.

3. JOSHUA D. GREENE AND AN FMRI INVESTIGATION OF EMOTIONAL ENGAGEMENT IN MORAL JUDGMENT

In *An fMRI Investigation of Emotional Engagement in Moral Judgment* by Joshua D. Greene and few others, they have made a variant on the Trolley Problem (I will call it Trolley Problem* for convenience) by pairing up **Bystander** and **Fatman** rather than **Transplant** [2]. Greene explicitly states in the beginning of the paper, because “there is no set of consistent, readily accessible moral principles that captures people’s intuitions concerning what behavior is or is not appropriate in these and similar cases” it causes the possible solution of this problem to stay obscure [2]. Hence, the authors decided to take a look at this problem through psychological experiments with the hypothesis of **Fatman** getting more emotional engagement in a way that **Bystander** does not, causing the tendency for different intuitive reactions.

To test the hypothesis, Greene used functional magnetic resonance imaging (fMRI) on the participants’ brains while providing them with “non-moral” (e.g. mathematical dilemmas), “moral-personal” (e.g. **Transplant** and **Fatman**), and “moral-impersonal” (e.g. **Bystander**) dilemmas to respond. As a result, the hypothesis was proven correct.

During the experiment, when “moral-personal” dilemmas were asked, brain areas associated with emotional responses had an “increased relative activation” and with the affirmative responses’ (e.g. agreeing to push in **Fatman**) reaction time was significantly longer than all other decisions made. Therefore, Greene concludes by stating that “emotional responses are likely to be the crucial differences between these two cases (**Bystander** and **Fatman**)”.

4. PETER SINGER AND ETHICS AND INTUITIONS

Peter Singer in his essay *Ethics and Intuitions* also tackles the Trolley Problem* but from a perspective of evolutionary theory. Assuming Greene’s conclusion of emotional differences is correct, Singer concludes that the emotional differences which cause our different intuitive outcomes originate far back in our ancestral backgrounds. He believes, as humans have been living in compacted communities that handled conflicts physically, “we have developed immediate, emotionally based responses to questions involving close, personal interactions with others” (e.g. moral-personal dilemmas); whereas switch related actions has not been deeply integrated (or has no relations at all) with our ancestral

life enough to provide us with a similar immediate emotional response [6].

However, although already forming a theory that seems normative, Singer rejects the necessity of having a normative moral principle. Unlike others who would have attempted to build upon the basis of the evolutionary theory which he has as a background, he simply rests on the thought of there being “little point in constructing a moral theory” that we already have evolved to have intuitive responses to [6].

5. FRANCES M. KAMM AND THE DOCTRINE OF TRIPLE EFFECT AND WHY A RATIONAL AGENT NEED NOT INTEND THE MEANS TO HIS END

The Doctrine of Triple Effect and Why a Rational Agent Need not Intend the Means to His End: Frances M. Kamm by Frances M. Kamm similarly looks upon The Trolley Problem* but through the lenses of intention, the Doctrine of Double Effect (DDE), and a later created Doctrine of Triple Effect (DTE). Kamm believes it is permissible to turn into a Bystander as hitting the one is not an intended action to save the five. In contrast, pushing and the fat man later being hit by the trolley in **Fatman** is a required and intended action, leading to no permissibility of this action. In addition, she believes it is permissible to turn the trolley in **Loop** because the hit on the one is also not an intended action.

To prove the permissibility of **Loop**, Kamm introduces a variant of **Loop** (we would call it **Loop***) which is similar to the original but the lever in this variant not only redirects the trolley but also pulls the one person and leads it to bump into the trolley. In **Loop***, she considers the act fully permissible. Although the death of the one is an effect of the agent pulling the lever, it would not be the intention of the act as the agent only intends to turn the trolley away from the five. To further explain, Kamm introduces another variant of **Loop** (we would call it **Loop****) which is also similar to the original but pulling the lever would release another trolley that would still kill the five. In **Loop**** Kamm similarly considers it as a permissible act. It is explained by the agent who still did not intend to kill the five. In this variant, there is no doubt that the original threat, the first trolley attempting to run over the five, has already been eliminated through the pull on the lever. It is also true that the act releases the second trolley which would still kill the five. Yet, the release of the second trolley would merely be creating another problem, similar to how it would be in **Loop** and **Bystander**, while it does not bother the intention of the act being diverting the upcoming threat (original trolley) for the five.

Unfortunately, some variants of **Loop** would actually be impermissible by DDE, since our intended act does not lead to the greater good (that the five be permanently

saved) and justifies the bad side effect [5]. However, proven by the logic previously stated, it should be morally permissible for us to turn the trolley into some variants of **Loop**. With this in mind, Kamm wishes to revise on DDE and change it into the Doctrine of Triple Effect (DTE):

“A greater good that we cause and whose expected existence is a condition of our action, but which we do not necessarily intend may justify a lesser evil and involvement leading to it that we must not intend but may have as a condition of action.” [5]

Kamm believes, with DTE, though it still lacks being looked upon, it would solve many cases which DDE would inappropriately judge the permissibility of while still having the supposedly impermissible situations, like pushing in **Fatman**, stay that way.

6. COMPARE AND CONTRAST

Within all theories previously introduced, the one difference that is the most obvious while being the most significant would be the difference in solving the trolley dilemma (and its variants). Judith Jarvis Thomson utilizes the widely adapted abstract notion, rights, with a comparatively more concrete notion of threat to form her “distributive exemption” theory. Joshua D. Greene and the other authors from their theory by adapting a psychological aspect with rigorous experiments and modern technology (fMRI) which lead to the formation of their final conclusion. Peter Singer includes the experiment results of Greene and builds upon it with his own evolutionary theory through an emphasis on our ancestors’ actions and reactions. Frances M. Kamm develops her theories from the fully subjective concept of “intention” while also involving the “outdated” DDE which was later replaced by her DTE.

Nevertheless, although the theories which attempted to solve this dilemma are from a variety of fields, would their difference also indicate a possibility of them being compatible with each other? I will attempt to test each out with the classic **Trolley** case.

In “distributive exemption”, an action must be acceptable if (i) it does not infringe anyone’s rights and (ii) it would redistribute a previously existing threat from the many to the few. Hence, I believe **Trolley** would be acceptable under this theory. The agent as the driver simply has two options, leaving the trolley as it is and turning the trolley onto the sidetrack, while in both cases the agent would not infringe anyone’s rights as he would merely be acting on the trolley. In addition, if the agent turns the trolley onto the sidetrack, he would redirect a preexisting threat from the many to the few, perfectly fulfilling the two requirements of Thomson’s theory.

In Greene’s paper, **Trolley** would be most likely classified as a “moral-impersonal” dilemma (similar to

Bystander) because it does not include direct contact with another person. In this sense, since **Bystander** has been stated as permissible since the beginning of the paper, I would believe that **Trolley** is similarly permissible.

According to Singer's evolutionary theory, **Trolley** would by no means be something that our ancestors would have encountered, turning a big steel machine towards another direction and causing a death, being similar to **Bystander**. Hence, **Trolley** would similarly have been a time too short for an emotional response to be passed down, and like **Bystander**, it would be permissible to hit the one.

Through the notion of "intend", the agent from the Trolley would surely be intending to turn the trolley away from the five while having the one being hit merely as a side effect. On top of that, DTE would also prove the permissibility of turning the trolley as this evil (hitting the one) would be justified by a greater good (saving the five).

Thus, it is very interesting that even when the four theories that possibly solve the trolley dilemma are from distinctly different aspects, they have still shown to be compatible with each other.

6.1. Compare and Contrast

A similarity that exists amongst the four theories is them being inevitably attempting to achieve a normative definition of morality (morality that would be accepted by all rational agents), wishing all who read their paper to view their theory as factual evidence [1]. It is true that in Greene's and Kamm's papers they all consider their final conclusion needing more revision. Yet, by the innate nature of essays, they have been inevitably trying to convince their readers that what they have done is correct and should be accepted. Even when they mention the necessity of future revision, it would still be possible for us to indicate that they believe their current theory/conclusion is a necessary basis for future developments.

However, as John Harris states, morality is "to make the world a better place and people better people" but different people may "have different ideas about what making the world a better place amounts to" [3]. In my understanding, as we are all individuals who constantly receive external blasts of information, it is inevitable that our interpretations of certain situations would differ from one another. It similarly adapts to the solutions of the trolley dilemma. Maybe for Thomson herself, she considers her "distributive exemption" theory the "ultimate solution" for all trolley dilemmas and it would certainly be permissible. Different people would inevitably have different "ultimate solutions" for the trolley dilemma. Due to this fact, maybe there would even be no normative "ultimate solution" that would ever

be found. Yet, this is exactly why dilemmas are invented. As theories crash, merge, and undermine each other, we are hoping to find the slightest fact which would be acceptable by all rational agents and even these would bring about enormous waves of better self-recognition as it builds a steady basis for the future.

7. CONCLUSION

I doubt that anyone who has read any of these papers would question the authors' professionalism on their respective theories. However, I would hope to point out the selfishness of humans that may possibly exist in the variants to which they are not directly responsible to (**Bystander** and **Fatman** are all examples with no direct responsibility as the agents are merely watching, with **Trolley** and **Transplant** being the contrary). Though altruism is an authentic existing notion, I suggest that it is only achieved while the action promises no essential interest of the actor itself is affected (e.g. not going into jail due to an action). Whereas when an essential interest is at risk, one would always prioritize themselves in any sort of action, like how one would be attempting to push themselves up when they are in the danger of drowning, even when the only push-able object around them is another live human being. Therefore, in variants that the agent is not directly responsible for, I would see another option that coexists alongside acting (e.g. pushing) and not acting (e.g. not pushing), which is to walk away. By letting the agent be completely outside of the intuition decisions, its advantage is obvious, which is to have no possibility of being morally judged. Nevertheless, if the option of walking away is available, I am not certain if this situation would be longer be categorized as a (variant of) trolley dilemma or not, as it takes the agent away from the "dilemma atmosphere".

No matter what variant of the trolley dilemma, just as I have mentioned previously, it would be with such significance if the trolley dilemma may be solved or even if all rational agents may stand on a common ground for an initial step of solving the dilemma. As it would not only bring an advancement on morality, it would also certainly affect fields such as euthanasia, abortion, automatic driving, legal judgment, and even on the more recent occasions such as the distribution of scarce medical resources on patients with COVID-19. Yet, before expecting the upcoming benefits, we should certainly attempt to address or find common ground on questions such as: are the solutions on the morality of the trolley dilemma normative (would be agreed by all rational agents) or descriptive (would be agreed only by a specific group of rational agents)? May theories that would be compatible with each other merge and form a better explanation of the trolley dilemma? The list of questions may go on, but if any of the questions may be later recognized or solved, I believe it would nonetheless be a leap on the process of solving the trolley dilemma.

REFERENCES

- [1] Gert, Bernard, and Joshua Gert. "The definition of morality." (2002).
- [2] Greene, Sommerville, Nystrom, Darley, and Cohen, "An fMRI investigation of emotional engagement in moral judgment," 2107.
- [3] Harris, John. "The Doctrine of the Triple Effect and Why a Rational Agent Need not Intend the Means to his End: John Harris: The Moral Difference Between Throwing a Trolley at a Person and Throwing a Person at a Trolley." In *Aristotelian Society Supplementary Volume*, vol. 74, no. 1, pp. 41-57. The Aristotelian Society, 2000.
- [4] Jarvis Thomson, Judith. "The trolley problem." *Yale Law Journal* 94, no. 6 (1985): 5.
- [5] Kamm, Frances M. "The Doctrine of the Triple Effect and Why a Rational Agent Need not Intend the Means to His End: Frances M. Kamm." In *Aristotelian Society Supplementary Volume*, vol. 74, no. 1, pp. 21-39. The Aristotelian Society, 2000.
- [6] Singer, Peter. "Ethics and intuitions." *The journal of ethics* 9, no. 3 (2005): 331-352.