2015

Frowe's Machine Cases

William Simkulet
Cleveland State University, w.simkulet@csuohio.edu

Frowe's Machine Cases

ARTICLE · SEPTEMBER 2015

READS
5

1 AUTHOR:

William Simkulet
Cleveland State University
13 PUBLICATIONS **2** CITATIONS

Frowe's Machine Cases

William Simkulet

Abstract
Helen Frowe (2006/2010) contends that there is a substantial moral difference between killing and letting die, arguing that in Michael Tooley's infamous machine case it is morally wrong to flip a coin to determine who lives or dies. Here I argue that Frowe fails to show that killing and letting die are morally inequivalent. However, I believe that she has succeeded in showing that it is wrong to press the button in Tooley's case, where pressing the button will change who lives and dies. I argue that because killing and letting die are morally equivalent we have no reason to press the button in the machine case. Pressing the button in this case is morally wrong because there is no reason to do it; to press the button is to treat matters of life and death irreverently.

Introduction

In Helen Frowe's "Killing John to Save Mary: A Defense of the Moral Distinction between Killing and Letting Die," she argues that there is a substantial moral difference between killing and letting die. She sets out to demonstrate the difference by analyzing Michael Tooley's machine case:

> *1. Machine* – Two children – John and Mary – have been placed inside two chambers in a machine. Between the two chambers is a canister of poison gas that will shortly be released into Mary's chamber. However, if a passerby presses a button on the machine, the gas will be released into John's chamber instead.[1]

[1] For the original version of this case, see Tooley (1980). For the purposes of each case discussed in this paper assume the agents involved are infallible about the consequences of their actions unless otherwise noted.

Tooley infamously contends that because killing and letting die are morally equivalent[2], the passerby ought to flip a coin to decide whether to press the button. Frowe argues that the passerby shouldn't press the button, nor should he flip a coin to decide whether to press the button because (1) there is a significant moral difference between killing and letting die (killing is worse) and (2) pressing the button would be a case of redirecting harm from one person to another – from Mary to John – and thus pressing the button is worse than doing nothing. Frowe argues that it is only acceptable to redistribute harm to others if they have what she calls "a "fair chance" to avoid being at risk of harm." (59) Frowe constructs a series of cases that she believes illustrates these two points.

This paper is divided into three sections. In the first, I look at Frowe's argument that killing and letting die are inequivalent. I argue that Frowe fails to show that there is a morally significant difference between killing and letting die. In the second, I look at Frowe's theory of redirecting harm, and argue that it is inconsistent with our commonsense moral intuitions regarding self-defense. Despite this, I contend that Frowe has given us the tools to show that pressing the button in the machine case is morally wrong regardless of whether killing and letting die are equivalent. In the third section I argue that if killing and letting die are morally equivalent, then the outcome of pressing the button is morally equivalent to the outcome of not pressing the button. As the outcomes are (by assumption) equivalent - either Mary dies or John dies - the agent lacks any moral reason to intervene. To press the button, then, is at least *prima facie* morally wrong because it is a waste of time and effort. Furthermore, there is another reason why we shouldn't press the button - or flip a coin - in this case - it treats matters of life and death irreverently.

I. On Killing and Letting Die

To illustrate the difference between killing and letting die, Frowe compares the following two cases:

> *2. Disease* – Both John^{D} and Mary^{D} have a fatal disease. Their doctor has a single dose of the antidote. "Neither John nor Mary has any prior claim upon the antidote." (Frowe, 57)

[2] To use James Rachels' terminology, all else being equal there are the same reasons against killing as there are letting die, and thus killing and letting die are morally equivalent. See Rachels (1975), (1979), (2001).

> *3. Diseased Mary* – Mary^{DM} has a fatal, non-communicable disease. She is trapped in a room with John^{DM} (sedated) and the instructions and materials necessary to make the cure – a gas poisonous to those without the disease. She will die before help arrives to free either of them from the room.

Frowe contends that in *Disease*, "one ought to toss a coin" to determine who lives; but in *Diseased Mary* "one should not toss a coin to see whether Mary can permissibly [kill] John." (57) There are two problems with comparing these cases to the machine case - (1) flipping the coin plays a different role in *Machine* than it does in *Disease*, and (2) *Diseased Mary* differs from *Disease* in that there is no impartial third party.

For Tooley, in *Machine*, the passerby finds himself caught in a Buridan's ass-type situation. He believes that all life is precious, and after examining the machine and trying to see if he can free both of the children, he concludes that there are only two options - (a) press the button (kill John, save Mary) or (b) not press the button (let Mary die, let John live). For Tooley, the coin flip serves as an *ad hoc* subjectively-indeterministic tie-break to free the passerby from his indecision. Tooley's passerby believes that both options are morally equivalent, and yet he has to choose between the two (otherwise he'd be stuck, unable to make a choice, like Buridan's Ass was said to be stuck indefinitely between two equally appetizing options, unable to choose between them).

In contrast, in *Disease*, for Frowe, flipping the coin is used as a fair and impartial means to determine who gets the antidote. John^{D} and Mary^{D} are equally deserving, but only one can get the antidote. Here the coin flip serves as an *ad hoc* subjectively-indeterministic tie-breaker, but unlike in Tooley's case, the coin flip is supposed to makes the outcome fair.

To illustrate this difference, suppose two passersby pass by Tooley's machine. The first finds herself equally drawn to pressing the button and not pressing it, and flips a coin to decide what she will do. Let's say she presses the button, redirecting the harm from Mary to John. The second passerby also flips a coin, then flips the switch redirecting the harm back from John to Mary. If the coin flip is an indeterministic means of breaking her mental stalemate, the first passerby might notice the second pressing the button - undoing her work - but she would have no reason to be offended or to go back and press the button again.

In contrast, suppose two doctors enter the room in *Disease*. For Frowe, the coin flip determines the just outcome. If a second doctor stopped the administration of the antidote and flipped a second coin, then gave it to the winner of the second coin flip rather than the first, this would be considered unfair and the first doctor would have a moral reason to intervene.

The second problem with Frowe's analysis turns on the fact that *Diseased Mary* is unlike the previous two cases in that there is no impartial third party. Coin-flips in both *Machine* and *Disease* were used by impartial third parties to guide their actions. Here Mary is not an impartial third party. The relevant question here is whether Mary is morally justified to kill in self-defense. Commonsense ethics, and many normative ethical theories, seem to hold that it is morally acceptable for innocent persons to kill other innocent persons in self-defense. Although such cases are rare, it is generally accepted that when all else is equal, we can put our own well-being ahead of that of others. Mary is justified in creating the cure that will incidentally kill John in *Diseased Mary* because she is acting in self-defense. For Frowe's purposes, though, we need a revised case:

> *3a. Diseased Mary in the Machine* – John^{DMM} and Mary^{DMM} – have been placed inside two chambers in a machine separated by a thin plastic wall. Mary^{DMM} has a fatal, non-communicable disease. In John^{DMM}'s chamber are all of the instructions and materials for making the cure – a gas fatal to those without the disease. By pressing a button on the machine, a passerby can dissolve the plastic wall between the two chambers.

Unlike *Diseased Mary*, this case is not a case of self-defense.

Frowe says "One way that we can explain the difference between *Disease* and *Diseased Mary* is by thinking about the courses of action that we could justify to John." (57) There are two substantive flaws with this stance. First, if we're interested in justifying our potential actions, we should be equally interested in justifying them to Mary as we are in justifying them to John.

Second, there is a substantive difference between justifying an act and justifying an act to John; the latter seems to imply that we need John's permission to morally engage in the act - in this case putting John in danger - but I suspect many wouldn't be inclined to consent to being put in danger, even if doing so was morally acceptable. For example, most people believe that killing in self-defense is morally acceptable - especially killing vicious

agents who freely put your life at risk. Suppose a vicious murderer kidnaps you and locks you in a cage deep in his basement and that, while readying his weapons, he explains that he is going to kill you. It occurs to you, however, that you might be able to save yourself if you kill him first. It would be quite absurd to suggest that you would need to be able to (counterfactually) justify your act of self-defense to the killer in order to kill in self-defense.

Most moral philosophers would have no trouble justifying the right to kill in self-defense, but it is a radically different question whether or not such a justification would be sufficient to justify it to the killer. Whether we can justify our action to the killer is irrelevant to whether or not we can kill the killer in self-defense.

In *Disease*, Frowe contends both John and Mary would consent to a coin flip deciding their fate; but that "In *Diseased Mary* John has no reason to agree to a third party's tossing a coin to decide whether Mary can manufacture the gas." (58) It's not clear that a third party has any bearing over Mary's actions in *Diseased Mary*, so for our purposes Frowe would contend that in *Diseased Mary in the Machine* John would have no reason to agree to the passerby's tossing a coin to decide whether to press the button. Still, this line of reasoning raises three problems.

First, it's not clear that Mary and John wouldn't advocate for some other *ad hoc* arbitrary decision making convention, such as "first-come, first served," the outcome of a game of checkers, etc. If we stipulate Mary and John are rational, self-interested individuals, then there's no reason to think they'd consent to a truly impartial decision making method at all - they'd prefer the method that would give them the best chance to live. This illustrates an important unparallel between Frowe's analysis of *Disease* and *Diseased Mary* - Mary's interests aren't consulted in the latter, and thus Frowe puts Mary at a disadvantage.

This unparallel is the second problem with Frowe's analysis. Here, Frowe seems to be begging the question - assuming that there is a morally relevant difference between killing and letting die, such that our actions need to be justified to John more so than Mary, because we'd be killing John, but merely letting Mary die. However, killing and letting die each affect the dying equally, so if we're interested in justifying our actions, we should be interested in what justification the passerby in *Diseased Mary in the Machine* could give to both parties to justify treating them unequally, to justify letting the person who put them in the machine decide their fate, or to justify letting luck (the flipping of a coin) decide who lives and who dies. Frowe contends

pressing the button cannot be justified to John, but certainly his inaction would be comparably uncomforting for Mary!

Third, Frowe's reliance on justification here seems inherently misguided - in some cases, John might agree to be killed to save Mary's life. Perhaps John is altruistic, perhaps he's suicidal. If John's consent matters here, this is morally relevant, such that Frowe should at least consider that in some cases killing might be preferable to letting die - when one has the consent of both parties.

II. On Redirecting Harm

Killing is wrong in *Machine*, Frowe contends, because it is redirecting harm to John "and John is not part of the lethal sequence of events that threatens Mary's life." (58) Frowe contends that it is not always wrong to redirect harms; she says doing so would be acceptable in cases like the following:

> *4. Body Armor* – Aggressor shoots at Victim. Bystander is nearby and can protect herself by putting on body armor, but refuses "because protective clothing is unflattering." (58-59) Victim can save himself only by deflecting a bullet towards Bystander.[3]

The relevant difference between *Body Armor* and *Machine*, she contends, is that Bystander had a prior chance to avoid even the risk of harm. Frowe contends "If the bystander had no chance to avoid her position… it is impermissible to kill [her] in… self-defense." (59) This suggests that if the bystander had a chance to avoid putting herself in harm's way and failed to do so (whether intentionally or negligently), then it may be acceptable to kill her in self-defense.

By the same token, if there is no morally relevant difference between killing and letting die, then in *Diseased Mary*, Mary^{DM} may be justified in acting to save her own life in such a way that will unintentionally kill John^{DM} if John^{DM} had previously had a chance to avoid being trapped in the room. Notably Frowe does not specify whether John^{DM} could have avoided being locked in the room with Mary^{DM}.

I agree that Victim would have different moral obligations towards a negligent Bystander than towards a virtuous Bystander; however her account in *Body Armor* is still radically inconsistent with our commonsense intuitions regarding killing in self-defense. Consider:

[3] For the purposes of this and the following case, bullets are always lethal.

> *4a. No Body Armor* – Aggressor^{A} shoots at Victim^{A} with his last bullet. Bystander^{A} is nearby – wrong place, wrong time. Victim^{A} can save himself only by deflecting a bullet towards Bystander^{A}.

While tragic, my intuition is that Victim^{A} is morally justified in acting in self-defense, even at the possible cost of an innocent person's life. (Of course Victim^{A} might, like John, justifiable choose to put the life of others ahead of his own, but this isn't required.) Self-defense cases are generally problematic because even those committed to the view that everyone's life is morally equivalent tend to have strong intuitions in favor of putting one's own life ahead of others in self-defense cases. One explanation is that (innocent) persons have a right to self-defense. Insofar as rights go, this one seems straightforward enough. Assuming moral agents have a right to life, in cases of scarcity and conflict there is a *prima facie* moral obligation to act to prevent conflict; but when there is no other course, killing in self-defense is *prima facie* morally acceptable.

However, because I am generally leery of rights-talk, an alternate explanation for why we are morally justified in choosing to preserve ourselves over others turns on our privileged access to our private mental states and moral history. For the moment, let's assume Victim^{A} is a generally good person. If this is the case, all else being equal, Victim^{A} has more reason to believe he is innocent than a stranger, and because he is morally obligated to favor innocent persons over villainous ones, he is obligated to favor himself over Bystander^{A}. This does not mean that Victim^{A} doesn't have any moral obligations to Bystander^{A}. If Victim^{A} has the option to either deflect the bullet and certainly kill Bystander^{A} or deflect it in such a way it would kill no one, all else being equal he is morally obligated to do the latter. Frowe contends that this kind of obligation would be lesser for "willing bystanders" in the same position as the one in *Body Armor* – but it would be quite odd if Frowe thinks that the killing would only be justified in terms of the relatively minor moral failing of being negligent. It's certainly not something that you can reliably justify to the negligent bystander, who - much like John - would probably vote against any action that would result in his death.

I think the privileged access account above is superior to the rights-based account because it explains our intuitions in rare cases where self-defense comes at a steep price. Consider the following case:

> *4b. Impending Nuclear Armageddon* – Terrorists have hacked into a nuclear armed submarine, and have aimed the missiles at a large number of highly populated targets. The only way to stop these missiles from launching is to activate the convenient new "self-destruct" system that responds to the captain's voice. As the captain begins uttering the self-destruct code, Ricky (a reporter covering the submarine) decides that he doesn't want to die and realizes that he can use his microphone cord to strangle the captain, preventing him to blowing up the ship, saving his life at the cost of millions of others.

According to the rights-based theory, it is morally acceptable (but not obligatory) for Ricky to kill the captain. However, according to the privileged access theory, because Ricky has overwhelming evidence that killing the captain will result in the deaths of many innocent persons, it is unacceptable to do so. Though, this is not to say that it is morally unacceptable to risk the lives of immoral persons to save yourself; consider this variation of *Body Armor*

> *4c. Willing Spectators* – Sparky has been enslaved and forced to fight in the Coliseum in front of legions of fight fans fully aware and apparently indifferent to the fact that he has been enslaved. One day, a lion lunges at Sparky, who has to choose whether to let the lion maw him, or to deflect the lion into the stands where he will no doubt kill many spectators.

Just as Frowe thinks it is justifiable to deflect a bullet in *Body Armor*, I think it is acceptable to deflect the lion in *Willing Spectators*. Even if the immoral actions of the spectators are not worthy of death, I think their immoral action absolves Sparky – and us – from having to worry about their well-being in such a case. They might not have deserved to be killed, but they did fail, morally. In contrast, in *Diseased Mary*, even if it is morally acceptable for Mary to kill John to save her own life, she shouldn't be happy about it.

I have the strong intuition it is acceptable to kill in *No Body Armor*, but Frowe argues it is wrong to deflect harm onto innocents. Still, I suspect she would conclude it is morally acceptable to kill innocents who are the sources of possible harm to you, for example virtuous soldiers on opposing sides during war. It strikes me as odd that it would be acceptable to kill innocent persons trying to fulfill their moral obligations, but not innocent persons in

the wrong place at the wrong time. However, Frowe contends that it is acceptable to deflect the bullet in *Body Armor* solely because of Bystander's moral failing, and contends it is never morally acceptable to kill "bystanders" (those who have not initiated a threat) in self-defense. Believing this, she turns to her penultimate case:

> *5. Armed Machine* – John^{AM} and Mary^{AM} – have been placed inside two chambers in a machine. Mary^{AM} will be killed unless a passerby presses the button. If the button is pressed, John^{AM} will be killed in her stead. Their kidnapper has armed John^{AM} and Mary^{AM} with modified automatic-weapons fixed to the outside of the machine. The trustworthy kidnapper tells them that one of the weapons is loaded with live ammunition, and one is loaded with blanks, and that the weapons can only fire on a warm heat signature of a human being.

Frowe contends that John^{AM} would be morally justified in firing on the passerby if he tried to press the button because his trying to press the button would be a threat. Although killing the passerby would get Mary^{AM} nothing, Frowe asks whether Mary^{AM} might be justified in shooting him in the knee and promising further force if he doesn't press the button. If Mary^{AM}'s action is justified, Frowe says, "we are committed to the implausible claim that one may use seriously harmful means to force a person to come to one's aid at the cost of an innocent person's [John^{AM}'s] life." (61) While regrettable, I don't find this implausible. Indeed, this is what I contend is acceptable in *No Body Armor;* that one is morally justified in acting in self-defense even at the cost of the life of an innocent person.

If Mary^{AM} is not justified in using this force, Frowe contends, it supports the view that there is a substantial difference between killing and letting die. (61) She goes on to say "… that you may not do *as much* against someone who refuses to save you as you may do against someone who is going to kill you is sufficient to support the killing/letting die distinction." This is just bizarre; consider a variation of *Armed Machine*:

> *5a. Solo Machine* – Mary^{SM} has been placed inside a machine that will release poison gas into her chamber when the clock hits zero unless someone presses a button. This is all very obvious to Mary and anyone who would pass her by. Her captors have given Mary a machine gun.

Suppose a numbers of passersby see Mary^{SM}'s plight, see that it is easy for them to save her life, and freely choose not to do so. Certainly Mary^{SM} is morally justified in both threatening to kill bystanders who would let her die, as well as following through with her threat if they fail to save her life - especially if doing so might make other bystanders press the button. The very notion that a passerby might witness Mary^{SM}'s plight and do nothing is morally abhorrent!

In *Solo Machine*, it seems that we *can* do as much against someone who refuses to save you as you may do against someone who is going to kill you - that is to say that there is no support for the killing/letting die distinction. The difference between *Solo Machine* and *Armed Machine* is that the passersby who freely let Mary die when saving her life would cost them next to nothing are uncontroversially moral monsters, while the passerby in *Armed Machine* certainly does not exhibit the same disregard for human life. If Mary is unjustified in acting in *Armed Machine*, surely it is because the character of her targets is different - they're not *clear* moral monsters for not killing an innocent person to save her life, where as in *Solo Machine* they would be clear moral monsters for letting her die for no reason.

III. Killing Arbitrarily

Frowe's final case, I think, has the most merit:

> *6. Blind Machine* – A passerby comes across a machine in which two children have been placed inside separate chambers. A canister of gas is hidden out of view above one of the chambers. Pressing the button will change the chamber it is aimed at.

There is no good reason to press the button here, as the passerby cannot tell who he will be killing and saving. *Blind Machine* illustrates the futility of flipping a coin in *Machine*. Because John and Mary are morally equivalent in *Machine*, there is nothing that the passerby can do that could bring about a morally different outcome. To press the button is to waste the passerby's effort, and perhaps to cause additional psychological harm to those trapped in the machine.

Blind Machine doesn't let Frowe help herself to the conclusion that killing is worse than letting die; pressing the button in *Blind Machine* is wrong because there is no reason to do so, and thus to press the button is to act irrationally and wastefully. Furthermore, pressing the button might cause

additional harm and suffering to those in the machine. Tooley's position isn't that killing and letting die are always morally equivalent, but that *all else being equal,* killing and letting die are morally equivalent. All *Blind Machine* has demonstrated is that *Machine* doesn't quite make everything else equivalent.

Suppose, though, that the passerby presses the button in *Blind Machine*. And suppose that she does it often because she likes having control over who lives and dies. This, it strikes me, is morally abhorrent because the passerby is making life or death decisions without regard to morality; her actions aren't done for moral reasons, rather they're done in wanton defiance of morality. This is a particularly egregious form of willful negligence, what we might call "playing God" because it involves carelessness with matters great importance - of life and death. I think it is uncontroversially true that this passerby acts immorally, and is severely morally blameworthy for her actions. Note that the actual or expected consequences of her actions are irrelevant to explaining what's wrong with her choice. Suppose that she presses the button twice - the gas momentarily switches targets, but then switches back. This has no effect on the outcome, but the passerby is clearly morally blameworthy for her careless attitude towards the life and death of others.

Conclusion

We are now in a better position to explain what is wrong with pushing the button in Tooley's machine case. If moral agents have (libertarian) free will, then they're not like Buridan's ass - they can make (arbitrary) choices between two outcomes on their own without flipping a coin. (The doctors in *Disease*, while capable of making arbitrary choices, probably should use an impartial, observable decision making method to demonstrate that their decision was arbitrary.) We can sensibly say that the passerby has at least three options - press the button, don't press the button, or flip a coin to decide whether to press the button. If killing and letting die are morally equivalent, all three options have a morally equivalent outcome and we can't decide between them by consequences alone. Because pressing the button and flipping the coin involve some effort, the default circumstance - not pressing the button is preferable. To decide for any other reason is to make a decision based on irrelevant grounds and to treat matters of life and death as if they didn't matter - and this is morally abhorrent.

References

Frowe, Helen. (2006/2010). Killing John to Save Mary: A Defense of the Moral Distinction between Killing and Letting Die. In *Action, Ethics, and Responsibility*, Joseph Keim Campbell, Michael O'Rourke, and Harry S. Silverstein (editors), The MIT Press: 47-66.

Rachels, James. (1975). Active and Passive Euthanasia. *The New England Journal of Medicine*, Vol. 292, pp. 78-80.

Rachels, James. (1979). Killing and Starving to Death. *Philosophy*, Vol. 54, No. 208, pp. 159-171.

Rachels, James. (2001). Killing and Letting Die. *Encyclopedia of Ethics, 2nd edition*, ed. Lawrence Becker and Charlotte Becker (New York Routledge, 2001), vol. 2, pp. 947-50.

Tooley, Michael. (1980). An Irrelevant Consideration Killing VS Letting Die. In *Killing and Letting Die*, Bonnie Steinbock (editor), Prentice-Hall: 56-62.

William Simkulet
Cleveland State University
simkuletwm@yahoo.com