



DISCUSSION NOTE

THOMSON'S TROLLEY SWITCH

BY ROBERT SHAVER

Thomson's Trolley Switch

Robert Shaver

The trolley problem is to justify thinking that it is (a) permissible to turn a trolley into one person to save five but (b) not permissible to push one person in the way of the trolley to save five. Judith Jarvis Thomson (2008) despairs of a solution. But, she argues, no solution is needed, since we were wrong to think it permissible to turn the trolley into the one. If so, our judgments in both cases can be explained by noting that killing is worse than letting die. I argue that Thomson is mistaken.

Thomson seems to argue as follows. Say I can

- (i) do nothing, letting five die
- (ii) turn the trolley into the one, killing the one and saving the five
- (iii) turn the trolley into myself, killing myself and saving the five.

I am permitted to not choose (iii); morality does not require that sacrifice. But if I do not choose (iii), I cannot choose (ii). I cannot make the one pay a cost I am unwilling to pay. If so, (i) is permitted.

Now consider the usual trolley case. I can

- (i) do nothing, letting five die
- (ii) turn the trolley into the one, killing the one and saving the five.

Thomson considers two possibilities.

A. Say that I would have been unwilling to turn the trolley into myself, had that option existed. Thomson comments of me that “[s]ince he wouldn’t himself pay the cost of his good deed if he could pay it, there is no way in which he can decently regard himself as entitled to make someone else pay it. .. Very few of us could decently regard ourselves as entitled to choose option (ii)” (2008: 366). If so, (ii) is not permitted.

B. Say that I would have been willing to turn the trolley into myself, had that option existed. If I turn the trolley into the one, I do something to the one to which he does not consent. I have a permissible alternative – (i). If so, (ii) is not permitted.

B is unconvincing. Thomson’s thought seems to be that since there is a downside to (ii) – I kill the one without his consent – and (i) is permitted, I should choose (i). But there is a downside to (i): five die. And as Thomson admits, most people think the downside of (i) is worse than the downside of (ii). Most people think this, despite knowing all along that the one is to be killed without his consent. This suggests that most people think that if (i) is permitted, (ii) is permitted.

B is also disappointing, since it implies that none of the new material concerning the three options and what I would have been unwilling to do

does any work. Thomson could have simply said that I cannot turn the trolley into the one because the one does not consent.¹

A is more interesting. The main problem with it is that one cannot infer from “I cannot decently regard myself as entitled to do x ” to “my doing x is not permitted.” Say I would not have been willing to pay some tiny cost, such as a blackened toenail, to save five. Thomson would infer that I cannot decently regard myself as entitled to blacken someone else’s toenail to save five. But even if I cannot decently regard myself as entitled to blacken someone else’s toenail, I am not only permitted but required to do so rather than let five die.

There are many cases in which I cannot decently regard myself as entitled to do something which I nonetheless am permitted (or even required) to do. Say that, when young, I avoided the draft. Now, as president, perhaps I cannot decently regard myself as entitled to enact the draft. But whether I am permitted (or required) to do so turns on the reasons supporting and opposing a draft – for example, is the curtailment of liberty worth the gain in military strength? – rather than on facts about me. If these reasons favor a draft, discovering that I cannot decently regard myself as entitled to enact the draft does not change what I ought to do. Similarly, I am required to blacken someone else’s toenail because the reasons for doing so (five are saved) outweigh the reasons against it.

We will not think well of my character if I choose (ii) when I would not have turned the trolley into myself. But people with flawed characters can perform right actions. In general (and as Thomson says elsewhere), the rightness of an action does not depend on the character of the person who performs it.²

In any case, concerns of character in the end support turning the trolley into the one. Say I am unwilling to turn the trolley into myself. My character would be better were I willing to do so. But since I am not, I am faced with the choice of inflicting a cost on the one I would not pay or letting the five die. Say I have more reason to save the five than to ensure that I do not inflict a cost on the one I would not pay. My character is made better by acting on the stronger reason than by acting on the weaker one, and so I should turn the trolley into the one.

¹ For similar objections to B, see FitzPatrick 2009: 642.

² Thomson 1999: 517. FitzPatrick has a different objection: Thomson’s case for thinking that I cannot kill the one if I can instead kill myself rests either on (a) appeals to our intuitions about other cases in which we think it is wrong to save by making someone else pay a cost we could have paid ourselves, or (b) appeal to a more general principle concerning respecting the one as a moral equal. He objects to (a) that this does not show anything about cases in which I make the one pay the cost as a side effect of diverting a threat. He objects to (b) that there is an alternative account of what counts as respecting the one which does not forbid me from killing him (2009: 638-40). But Thomson could reply that her argument rests neither on analogies to non-Trolley cases nor appeal to a more general principle. Her argument rests on our agreement that in Trolley cases I cannot kill the one when I could instead kill myself.

Thomson might object that in the trolley case, I am permitted not to bear the cost. In the toenail case, I am required to bear the cost; in the draft case, I have (by hypothesis) done something wrong. But it is not clear that this matters. Thomson's inference seems to be from what I can decently regard myself as entitled to do to what is not permitted. If I came to believe that in the trolley case I am required to bear the cost but would not pay it if I could, it seems Thomson would again say that I cannot decently regard myself as entitled to make the one bear the cost, and so am not permitted to do so. (It is also not clear that Thomson thinks I am required to bear the cost of the toenail.)

Thomson might also object that my unwillingness to turn the trolley into myself is itself a reason for me not to turn the trolley into the one. If asked why I do not turn it into the one, I might say, "Because I wouldn't turn it into myself." Even if this is true, however, she needs this reason against turning to defeat the reason in favor of turning, namely that five will be saved. Ensuring that I am in some sense consistent is less important than saving the five. (Perhaps "Because I wouldn't turn it into myself" is not the reason, but rather points to the reason: I wouldn't turn it into myself because I do not consent to dying, so I shouldn't turn it into the one since he does not consent either. This returns the argument to B.)

Another problem is that even if Thomson is correct in A that I cannot turn the trolley into the one, it does not follow that I cannot take steps which result in the trolley being turned into the one. I might say to myself, "Given that I am unwilling to send the trolley into myself, I cannot send the trolley into the one. I should disqualify myself from deciding, and find someone who can act simply on the basis of the reasons for and against turning." (The president might find the analogous line of thought compelling.) Say I find someone else to decide, and she would have been willing to turn the trolley into herself, had that option existed. If the argument of B fails, she is permitted to turn the trolley into the one.

For Thomson, the three-option case merely sets up the argument for the two-option case. But it raises curious issues of its own. It is plausible to rank the acts, from best to worst, as

I turn the trolley into myself
I turn the trolley into the one
I do nothing.

If I am permitted not to bring about the best outcome, and I act on this permission, one would think that I should now choose the next best outcome.³

³ A referee noted that there are alternatives to simply choosing the next best outcome. The referee endorsed flipping a coin between myself and the one. Whether or not one agrees that redirection requires this, Thomson at least must argue that adopting some apparently fair procedure is more objectionable than doing nothing. She does reject a coin toss in the three-

Thomson avoids this by supposing that if I do not turn the trolley into myself, I cannot turn it into the one. She concludes that not turning is permitted. But this assumes that the objection to my turning the trolley into the one is more serious than the objection to not turning. The objections to my turning it into the one are that the one dies and I should not make the one pay a cost I am unwilling to pay. If I do, I have a poor character. But the objection to doing nothing is that five die. That is a more serious objection.

One way to make the point is to suppose that I do not turn the trolley into myself because I am too scared to do so. If I then turn the trolley into the one, that might seem to increase the stain on my character. But if I do nothing, that increases the stain far more. I should make the best of a bad situation and choose the second best act. If so, even in a three-option case in which I am unwilling to turn the trolley into myself, I am permitted to turn it into the one.

Before closing, consider a different argument, turning on

- (P) I am permitted to make the one bear the cost if I would have been required to bear the cost, but am not permitted to make the one bear the cost if my bearing the cost would have been supererogatory.

What I am willing to do is irrelevant. This could be offered as a friendly recasting of Thomson or as an interpretation of her argument.⁴ (P) permits action in the toenail case but not in the trolley case (supposing that I am required to bear the blackened toenail). As an interpretation, this seems to make a great deal of her paper irrelevant to her real argument. She often does not treat willings as relevant only to character, and not to permissibility (2008: 366, 369, 370). However, in a two-option case in which I am the driver, and so will be killing one or five, Thomson thinks that even if I would have been unwilling to sacrifice myself, had the option existed, I not only can but must turn the trolley into the one (2008: 372). Here willings *are* irrelevant. So, interpretation aside, consider (P).⁵

option case on the grounds that it “is not fair [of the bystander] to impose a 0.5 risk of death on a person in order to do what it is morally optional for him not to do.” In a three-option case in which I am the (faultless) driver, she leaves open the possibility that I am permitted to toss a coin between myself and the one (2008: 371). More needs to be said about the connection, if any, between fairness and whether an act is supererogatory to know who is right. For example, one might think fairness requires only treating equals equally. That thought seems to be what justifies thinking that when I am the faultless driver, I am permitted to flip a coin. This might favor thinking that there is a *better* case for a coin toss between bystanders, in the original three-option case, than between the driver and bystander.

⁴ Max Wolinsky gave it to me as the former. A referee suggested it as the latter, on the plausible ground that Thomson would not let considerations of what one is willing to do bear on what is permissible, just as, in her criticism of double effect, she does not let considerations of what one intends bear on what is permissible (Thomson 1999: 515-17).

⁵ Perhaps Thomson’s view is that willings are irrelevant to permissibility in cases like the two-option driver, where we are convinced, independent of willings, that one is required to kill the one, but relevant to permissibility in cases like the two- and three-option bystander.

One worry is that we have more confidence in our judgment that we are permitted to turn the trolley than we have in (P). The trolley case seems a problem for (P) just as it is for the principle that killing is worse than letting die.⁶ Another worry concerns why the fact that my bearing the cost would have been supererogatory makes it the case that I am not permitted to make the one bear the cost. Say, as is plausible, that my bearing the cost is supererogatory because an agent-centered prerogative is justified, and not outweighed by saving the five. This would not explain why I am not permitted to make the one bear the cost.⁷

Robert Shaver
University of Manitoba
Department of Philosophy
bshaver@cc.umanitoba.ca

References

- FitzPatrick, W. J. (2009). "Thomson's Turnabout on the Trolley." *Analysis* 69.
Thomson, J. J. (1999). "Physician-Assisted Suicide: Two Moral Arguments." *Ethics* 109.
____ (2008). "Turning the Trolley." *Philosophy and Public Affairs* 36.

But this would need to be justified. And many are convinced, independent of willings, that one is permitted to kill the one in the two-option bystander case.

⁶ This is similar to FitzPatrick's objection to (a) in note 2 above.

⁷ Thanks to Joyce Jenkins, Jeff Verman, Sandy Vettese, members of my 2010-11 ethics seminar and moral philosophy classes, and an anonymous referee.