Boulders and Trolleys

D. W. HASLETT
University of Delaware

This discussion attempts to show that the elusive solution to the trolley problem lies hidden in the solution to another perennial problem in moral philosophy: the ducking puzzle. The key to solving the ducking puzzle is an important, but overlooked, exception to our obligation not to harm others, an exception for 'fair competition', which, it is argued here, is also the key to solving the trolley problem.

The 'trolley problem', originally formulated by Philippa Foot,[1] and developed further by Judith Jarvis Thomson,[2] is one of the more perplexing problems in moral philosophy. It is also among the more important, raising, as it does, serious questions about when it is morally permissible to kill people. An equally perplexing and important problem in moral philosophy – a problem that may appear unrelated to the trolley problem – is known as the 'ducking puzzle', formulated by Christopher Boorse and Roy Sorensen.[3] One proposed solution to the ducking puzzle is built upon a theoretical framework that includes an exception for 'fair competition' to our obligation not to harm others, an exception like that for self-defense.[4] But the exception for fair competition is so ubiquitous that, like the air around us, we pay little attention to it. Moral philosophers normally overlook it entirely. Yet, as this discussion will try to show, overlooking it is a mistake, since this exception is the missing key for solving not only the ducking puzzle, but the trolley problem as well.[5]

I wish to thank Marjorie Parsons for her helpful comments on this article.

[1] Foot formulates the problem and proposes her solution in Philippa Foot, 'The Problem of Abortion and the Doctrine of Double Effect', *Oxford Review* 5 (1967), pp. 5–15; reprinted, among other places, in *Ethics: Problems and Principles*, ed. John Martin Fischer and Mark Ravizza (Orlando, FL, 1992), pp. 60–7. The page reference here to this essay is from Fischer and Ravizza, *Ethics*.

[2] Judith Jarvis Thomson, 'The Trolley Problem', *The Yale Law Journal* (1985), pp. 1395–1415, reprinted in Fischer and Ravizza, *Ethics*, pp. 279–92, from which the page references here to this essay come. See also an earlier article by Thomson entitled 'Killing, Letting Die, and the Trolley Problem', originally printed in *The Monist* (1976), pp. 204–17, and reprinted in Fischer and Ravizza, *Ethics*, pp. 69–77.

[3] Christopher Boorse and Roy A. Sorensen, 'Ducking Harm', *The Journal of Philosophy* 85 (1988), pp. 115–34; reprinted in Fisher and Ravizza, *Ethics*, pp. 77–91.

[4] A preliminary version of this solution was set out in D. W. Haslett, 'Murder and the Exception for Fair Competition', *Social Theory and Practice* 29 (2003), pp. 631–54. The final version is set out here.

[5] A good sample of the many discussions of the trolley problem have been printed, or reprinted, in Fischer and Ravizza, *Ethics*.

© Cambridge University Press 2011
Utilitas Vol. 23, No. 3, September 2011
doi:10.1017/S0953820811000148

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

In any paper on moral philosophy not every belief about morality that is relevant to the argument can, of course, be defended within the space available; certain beliefs must be presupposed. As long as these presuppositions do not cross the line from being merely controversial to what many would consider absurd, they may, for purposes of the discussion, reasonably remain undefended. But it will be useful if certain undefended presuppositions relevant to this discussion are made explicit from the start. First, it is presupposed here that the most justified code of morality is a deontological code, one in which an act can be morally wrong independent of its causing any harm. Among the ways in which an act can be morally wrong – violate a moral obligation – independent of its causing any harm are by its (1) being deceitful; (2) deliberately causing another to be physically moved involuntarily, or prevented from moving; and (3) violating a promise or other morally binding commitment. Of course moral obligations have exceptions 'built' into them, such as for self-defense, or for permission. But more on exceptions later. It is also presupposed here that, although feeding the starving, rescuing strangers from impending dangers, and otherwise aiding others can be highly praiseworthy, strangers do not have a moral *right* to our aid; accordingly, we are not morally *obligated* to give it. Aiding them is, rather than being obligatory, a matter of good values.[6]

Finally, except when stated otherwise it is presupposed here that, in the cases to be discussed, none of those facing a danger are – because of who they are, or how they got into danger – any more, or less, deserving of aid than any others facing the same danger.

I. THE TROLLEY PROBLEM

The original trolley problem arises from two hypothetical cases, the first of which is as follows.

Trolley Driver: The breaks have failed on a trolley that is hurtling toward five people who are stuck on the track ahead of it, all of whom will be killed if the trolley collides with them. On a spur track going off to the right, one person is stuck. The trolley driver throws a switch that diverts the runaway trolley onto the spur track, thereby saving the lives of five at the cost of killing only one.

[6] More precisely, according to this presupposition we have no 'categorical' obligation to help those in need, as opposed to an obligation to help that we have voluntarily undertaken through a special commitment or relationship (a 'commitment' obligation), or through some wrong we have committed (a 'restitutional' obligation). In defense of this presupposition, see D. W. Haslett, 'Values, Obligations, and Saving Lives', *Morality, Rules and Consequences*, ed. Brad Hooker, Elinor Mason and Dale E. Miller (Edinburgh, 2000), pp. 71–104.

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

Most people's moral intuitions tell them that diverting the runaway trolley onto the spur track is, if not mandatory, at least morally permissible.

Now compare *Trolley Driver* with the second hypothetical case.

Transplant: Allie, who greatly sympathizes with hospital patients in need of transplants, kills Bailey, an innocent visitor at the hospital, so that Bailey's organs can be transplanted by a willing surgeon into five patients who, without these transplants, will surely die. Allie thereby saves the lives of five at the cost of killing only one.

We are to assume here that Allie qualifies as a bystander. For a person, *A*, to qualify as a bystander in cases in which people are threatened with harm from a danger, *A* cannot herself be one of the people facing the danger. *A* must also have no relationships that create a special obligation of care to any of those facing the danger. In other words, none of those facing the danger can be *A*'s child, patient, promisee, or anyone else to whom *A* has, for some reason, a special, morally binding commitment of care. So we are to assume that, to Allie, all of the patients are strangers.

Our moral intuitions tell us that for Allie to kill Bailey against his will to save the lives of these five patients is morally wrong, very wrong in fact. But then how can our intuitions about *Trolley Driver* be reconciled with our intuitions about *Transplant*? It appears that, in both cases, one person is killed to save five. What then, if anything, is the morally relevant difference between these two cases? This is the original trolley problem.

Not much shall be added here to the solution that Foot proposes. Referring to obligations as 'duties', Foot writes, 'The steering driver faces a conflict of negative duties, since it is his duty to avoid [killing] five and also his duty to avoid [killing] one.'[7] By deliberately diverting the trolley onto the spur track so that it runs over the person stuck there, clearly the driver kills one person. But if instead the driver had done absolutely nothing, not moved even a single muscle, thereby letting the trolley continue on its path toward the five stuck on the main track, this would have counted as his 'killing' five people. To see why, consider the following case.

Automobile Driver: Anthony is driving along a straight road in his car. Far ahead of him he sees Bethany directly in his path. She has fallen down, hit her head, and is lying in the road unconscious. Anthony can easily steer around her, but, if he does not move a muscle, he will hit and kill her. Not moving a muscle, Anthony deliberately keeps going straight ahead toward Bethany, and kills her.

[7] Foot, 'The Problem of Abortion', p. 65.

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

Anthony, in control of a car, is morally responsible for its movement. Since he could prevent the car from running into Bethany, but deliberately chooses not to do so, Anthony is therefore morally liable for causing Bethany's death; causing her death violates the moral obligation not to kill innocent people. Similarly, the trolley driver, in control of the trolley, is able to prevent it from killing the five innocent people stuck on the main track by diverting the trolley onto the spur track away from them. Had he deliberately chosen not to do so, he would have killed them, just as Anthony kills Bethany. So whether the trolley driver diverts the trolley onto the spur track or not, *either way* he will kill one or more innocent people, contrary to what the obligation not to kill innocent people normally prohibits. Thus he is faced with conflicting obligations. There is, however, built into every obligation this exception: when faced with conflicting obligations that arise through no fault of the agent, as long as the agent adheres to the obligation with priority, the agent is justified in doing what the other obligation normally prohibits.

In the case of *Trolley Driver*, which of the conflicting obligations, if either, has priority? A principle of morality to which we can appeal for answering this question is one something like this: From among alternatives that otherwise are morally equal, choose the one that causes, or is likely to cause, the least harm. Call this the 'least-harm' principle. Notice that the least-harm principle requires only that we choose the alternative that *causes* the least harm; it does not require that we *prevent* any harm, as admirable as preventing harm may be. And 'harming' others refers here to damaging their bodies, property, or financial position. In the case of *Trolley Driver*, from among the two otherwise equally prohibited alternatives, the one that causes the least harm is the alternative that kills one person, not five. Thus, the least-harm principle requires that the driver choose that alternative. So it is the obligation not to kill five that has priority, and the exception for conflicting obligations therefore justifies his doing what the obligation not to kill one normally prohibits.

How then does *Trolley Driver* differ from *Transplant*? Allie, by killing Bailey to save the five patients, violates the obligation to not kill innocent people. But for Allie, a mere bystander, saving the five patients is not obligatory. So the morally relevant difference between *Transplant* and *Trolley Driver* is this. Contrary to *Trolley Driver*, in *Transplant* the agent is not faced with *conflicting* obligations; in *Transplant*, Allie's obligation not to kill Bailey stands unopposed.[8]

[8] According to how the case of *Transplant* is usually formulated, it is not a bystander who kills Bailey; rather it is the surgeon in charge of the five patients who, in the best interests of her patients, kills him. Had it been the surgeon who killed Bailey, her

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

But, with the help of the following case,[9] Judith Jarvis Thomson presents us with a new version of the trolley problem, one to which Foot's solution does not apply.

Bystander at the Switch: This case is the same as *Trolley Driver* except that, once the driver of the runaway trolley realizes that its brakes have failed, he faints. But a bystander, waiting for a later trolley, is standing near a switch that can divert the runaway trolley onto the spur track. Knowing that a person is stuck on the spur track, the bystander nevertheless throws this switch, thereby saving the lives of five at the cost of killing just one.

As Thomson observes, most people's moral intuitions tell them that it is as permissible for the bystander to throw the switch as it would have been for the trolley's driver to have done so. Yet, as Thomson points out, if the bystander had *not* thrown the switch, it is plain that the bystander would not have *made* the trolley run into the five killing them; the bystander would merely have *let* it run into them.[10] In other words, if the bystander had not thrown the switch, he would not have *killed* the five, he would merely have *refrained from saving them.* And, as is presupposed here, even though saving strangers normally exemplifies good values, we have no *obligation* to do so. Indeed, the bystander is no more obligated to save the five than is Allie, in *Transplant*, obligated to save the five patients in need of transplants. So, just like Allie, the bystander is not faced with *conflicting* obligations. The bystander's obligation not to kill the one, like Allie's obligation not to kill Bailey, stands unopposed. Thus we cannot distinguish *Bystander at the Switch* from *Transplant* in the same way as we distinguished *Trolley Driver* from *Transplant*: namely, by appealing to conflicting obligations along with the least-harm principle. So how then can we distinguish *Bystander at the Switch* from *Transplant*? What exactly is the morally relevant difference between these cases? This is Thomson's version of the trolley problem, the version to which the remainder of this discussion is devoted.

II. THE DUCKING PUZZLE

To understand the solution to Thomson's version of the trolley problem, we must, it is argued here, first understand the solution to the

obligation not to kill – a 'categorical' obligation – would not have stood unopposed; it would have been in conflict with the Surgeon's 'commitment' obligation to act is in the best interests of her patients. But, in a conflict between a categorical and a commitment obligation, morality requires that we give priority to the categorical obligation. So by killing Bailey, the surgeon would have been just as guilty of wrongful homicide as was Allie. This is explained fully in D. W. Haslett, 'Conflicts and Commitment Obligations', *Public Affairs Quarterly* 18 (2004), pp. 319–36.

[9] Thomson, 'The Trolley Problem', p. 281.

[10] Thomson, 'The Trolley Problem', p. 282.

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

ducking puzzle. The ducking puzzle, which Christopher Boorse and Roy Sorensen set out, but did not claim to have solved, goes like this.[11] Consider, first, the following pair of cases.

(1) *Boulder 1*: *A* and *B*, strangers to one another, are standing in line on a mountainside to purchase a ticket to ride farther up the mountain to witness the view. The line faces down the mountain, with *A* the last person in the line, and *B* the next to last. A huge boulder becomes dislodged, and is hurtling toward *A* with just enough velocity to kill one person. *A* notices the oncoming boulder just in time to jump safely out of its way. The boulder then hits and kills *B* instead of *A*, as *A* knew it would.

(2) *Boulder 2*: This case is the same as *Boulder 1* except that, rather than jumping out of the way, *A* quickly grabs *B*, lifts her up and moves her between himself and the oncoming boulder so that she shields him from it. The boulder then hits and kills *B* instead of *A*, as *A* knew it would.

Since, in *Boulder 2*, *A* deliberately moves *B* into the path of the oncoming boulder, we view *A* as wrongfully killing *B*, just as if *A* had deliberately pushed *B* into the path of an oncoming truck. Yet, in *Boulder 1*, where *B* is killed because *A* 'ducks' the harm by jumping safely out of the way, we do not view *A* as doing anything morally wrong. But, in both cases, *A*, for self-preservation, avoids the danger from the oncoming boulder by shifting the danger onto *B*, knowing that, as a result, *B* will be killed. And, in both cases, what *A* does to shift the danger has the same result – *B*'s death – a result that, in each case, is equally foreseeable. So why is *A* morally liable for *B*'s death in *Boulder 2*, but not in *Boulder 1*? What exactly is the morally relevant difference between the two cases? This is the ducking puzzle.

Let us say that *A* is 'morally liable' for any foreseeable harm that *A* causes if and only if *A*'s causing the harm violates a moral obligation. And, for most practical purposes in law and morality, person *A* is said to 'cause' result, *R*, if and only if what *A* does is a necessary component of a series of events that are jointly sufficient for the occurrence of *R*. Or, for short, we may say that *A* 'causes' *R* if *R* would not occur except for what *A* does. In the 'boulder' cases, the 'danger' is a danger from something that has the potential to cause harm, as does an oncoming boulder, a runaway trolley, a lighted firecracker, a falling bomb, and a lethal virus. A danger *from* something is to be distinguished from a danger *of* something happening, such as dying, losing money, breaking a bone, making a bad mistake, or becoming addicted to drugs. Dangers

[11] Boorse and Sorensen, 'Ducking Harm'.

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

'from' refer to potential *causes* of what is bad, while dangers 'of' refer to potential *results* that are bad. The ducking puzzle is about dangers in the former sense.

Consider now a variation of the ducking puzzle exemplified by the following two cases, ones which suggest how common cases exemplifying this puzzle actually are.

Widget 1: *A* and *B* are competitors in the business of manufacturing and selling widgets. *A* takes a considerable amount of business away from *B* by making a better widget.

Widget 2: This case is the same as *Widget 1*, except that *A* instead takes the business away from *B* by spreading a rumor that *B*'s widgets cause cancer, even though *A* knows they do not.

In the two widget cases, it is more natural to characterize what *A* does not as shifting a danger away from himself onto *B* as in the two boulder cases, but as shifting, partially at least, an opportunity away from *B* onto himself, an opportunity for making money by manufacturing and selling widgets. Nevertheless, the puzzles exemplified by these two pairs of cases are variations of essentially the same puzzle: Why, in the second case within each pair, is *A* morally liable for the harm he knew that he would cause *B*, whereas, in the first case, he is not?

III. A SOLUTION TO THE DUCKING PUZZLE

The solution to that variation of the puzzle exemplified by the widget cases is, rather obviously, this: What *A* does in the first widget case to increase his opportunity for selling widgets while decreasing *B*'s is to make a better product, and making a better product is not morally wrong. So, *independent* of the harm it does to *B*, there is nothing morally wrong with what *A* does. But what *A* does in the second widget case to increase his opportunity at the expense of *B*'s opportunity – falsely proclaiming that *B*'s widgets cause cancer – is indeed morally wrong *independent* of the harm it does to *B*. Deliberate defamation by making false claims is morally wrong in itself, independent of any harm it may, or may not, cause others. This then is why *A* is morally liable for the harm to *B* he causes in the second widget case, but not in the first. Essentially this same solution can, it is claimed here, be applied to variations of the ducking puzzle in which *A*, instead of increasing an *opportunity* for himself at the expense of decreasing it for *B*, decreases a *danger* to himself at the expense of increasing it for *B*.

To understand this solution, it is necessary to understand exactly what is meant here by 'independently wrong'. To be independently wrong, what *A* does to shift a danger, or opportunity, must meet two conditions. First, it must, all things considered, be morally wrong;

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

that is, contrary to a moral *obligation*. There cannot be any excusing or justifying conditions. Second, what *A* does to shift the danger, or opportunity, must be not just morally wrong, but morally wrong for a reason that does not depend upon anyone's being harmed, or likely to be harmed, from the very danger, or from losing the very opportunity, that is thereby shifted. It must, conceptually, constitute a distinct wrongdoing.

Consider, for example, *Boulder 2*, where what *A* does to shift the danger from the oncoming boulder away from himself onto *B* is to pick *B* up and move her between himself and the oncoming boulder so that she shields him from it. The moral wrongness of *A*'s doing this does not depend upon anyone being harmed, or likely to be harmed, by the oncoming boulder, the very danger that *A* is shifting. Picking people up and moving them without permission violates a moral obligation by itself, regardless of any harm that may result from doing so. Thus what *A* does to shift the danger onto *B* is *independently* wrong. And since it is independently wrong, *A* is, according to this solution to the ducking puzzle, morally liable for the foreseeable consequences to *B*. One foreseeable consequence is *B*'s death. Therefore, *A* is morally liable for *B*'s death.

Consider, on the other hand, *Boulder 1*. In this case, *A* shifts the danger from the oncoming boulder away from himself onto *B* simply by jumping away. If *A*'s jumping away were morally wrong at all, it could be morally wrong only because, as a result, *B* would be harmed, or likely to be harmed, by the oncoming boulder, the very danger that *A*, by jumping away, is shifting. But if this is the only reason it could be morally wrong, then, by definition, this disqualifies it from being *independently* wrong. Finally, if what *A* does to shift the danger does not qualify as being independently wrong, then, according to this solution to the ducking puzzle, *A* is not morally liable for any harm to *B* from this danger.

There can also be cases in which, without his doing anything that is independently wrong, *A* shifts a danger away from himself onto *B* not by moving himself away from the danger, but by moving the danger away from himself. Consider, for example, the following case, a slightly simplified variation of an actual case in law.[12]

Firecracker 1: A lighted firecracker is thrown into a marketplace so densely crowded that it is bound to land next to someone. It lands next to *A* who, hemmed in by other people, is unable to run from it. To protect himself, *A* randomly tosses the firecracker away, whereupon it is bound to land next to someone else. It lands next to *B* and explodes, putting *B*'s eye out.

[12] *Scott v. Sheperd*, 2 Wm. Blackstone 893, 96 Eng. Rep. 525 (1773).

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

In the actual case, the court held that *A* was not legally liable for the harm to *B*, that the only person liable was the person who lit the firecracker and threw it into the crowd in the first place. Moreover, it is claimed here, *A* is not *morally* liable either. In this case, *A* shifts the danger from the lighted firecracker away from himself onto *B* simply by tossing it away. If his tossing the firecracker away were morally wrong at all, it could be morally wrong only because, as a result, someone would be harmed, or be likely to be harmed, by the lighted firecracker, the very danger that *A* is shifting. But if the only reason it could be morally wrong thus depends on someone's being harmed, or likely to be harmed, *by the very danger that* A *is shifting*, then, once again, this disqualifies what *A* does to shift the danger from being *independently* wrong. And if what *A* does to shift the danger does not qualify as being independently wrong, then, according to this solution to the ducking puzzle, *A* is not morally liable for any of the consequences of his shifting the danger.

Notice how *Firecracker 1* differs from a case in which it was *A* himself who lit the firecracker that he tossed at *B*, thereby putting *B*'s eye out. Call this case *Firecracker 2*. In *Firecracker 2*, as opposed to *Firecracker 1*, *A* is morally liable for the harm that *B* suffers. Yet, in both firecracker cases, the mere tossing away of the firecracker, independent of any harm likely to result from doing so, is not morally wrong. So, in neither case, is tossing it away *independently* wrong. Why then, in *Firecracker 2*, is *A* morally liable for the harm likely to result, but not in *Firecracker 1*? Here is why. One is always morally liable for harm resulting from a serious danger that, unjustifiably and inexcusably, one *originates*, regardless of whether or not originating it involves doing anything independently wrong. In *Firecracker 2*, but not in *Firecracker 1*, *A* originates the danger.

To help avoid any misunderstanding, perhaps one more 'danger' case, *Automobile Driver 2*, is worth considering here.

Automobile Driver 2: This case is the same as *Automobile Driver*, except for these differences. First, rather than driving his car on a wide, straight road with perfect visibility, Anthony is driving on a curvy, one-lane mountainous road, so narrow his car barely fits on it. Second, at an especially sharp curve, Anthony, through no fault of his own, comes suddenly upon Bethany lying there unconscious, leaving Anthony no way to avoid hitting and killing her other than by quickly swerving off the mountain, plunging to his death. Anthony does not swerve off the mountain, but instead hits and kills Bethany.

In this case, Anthony did, to be sure, originate the movement of the car, which eventually became a serious danger to Bethany. But since Anthony had no reason to *foresee* that the car's movement would become a serious danger to anyone, his having originated this danger is

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

excusable, and, being excusable, thereby relieves Anthony from moral liability for the harm to Bethany resulting from the danger.

A solution to the ducking puzzle is now close at hand. Two pieces to the solution are, however, still missing. For the first piece, consider the following variation of *Boulder 1*.

Model Airplane: This case is the same a *Boulder 1* except as follows. *B*, a model-airplane enthusiast, brought with her to the mountain a model airplane which she planned to display at a nearby model-airplane show. She set it down a few feet away from where she and *A* were standing in line, so she would not have to carry it while waiting to buy her ticket. The model airplane happened to be positioned in just such a way that *A* knew that, as he jumped away from the oncoming boulder, he had no choice but to land on the model airplane, destroying it.

The willful destruction of another's property without permission is normally morally wrong. And, by jumping away from the boulder, *A* does willfully destroy *B*'s property. So, for the sake of argument, let us assume that, in shifting the danger from the oncoming boulder by jumping away, *A* does indeed do something morally wrong. But if wrong, it would be wrong for a reason that is altogether independent of the harm that results from the danger. Thus, what *A* does to shift the danger would have to count as *independently* wrong, just as it does in *Boulder 2*, where *A* is morally liable for *B*'s death. Yet *Model Airplane* actually seems much closer to *Boulder 1* than to *Boulder 2* in that, just as in *Boulder 1*, *A*, in *Model Airplane*, shifts the danger away from himself merely by jumping away. Jumping away does not make *A* morally liable for *B*'s death in *Boulder 1*. Therefore, it should not make *A* morally liable for *B*'s death in *Model Airplane* either, even if he is morally liable for destroying *B's* model. So the problem is this: Why does independent wrongness make *A* liable for *B*'s death in *Boulder 2*, but not in *Model Airplane*?

The reason is that *A* should be morally responsible for *B's* death only if the independent wrongness of what A does is actually the *means* by which he avoids the danger, not a mere *side effect*. Although, in *Model Airplanes*, we are assuming that *A's* destruction of the model is wrong, independently wrong in fact, A's destruction of the model is merely a side effect of his shifting the danger away from himself onto *B*. The actual means by which he shifts the danger is by jumping away. Jumping away would, obviously, have shifted the danger away from himself even if he had not landed on *B*'s model. So even if what *A* did in *Model Airplane* were indeed independently wrong, he should not be liable for *B's* death any more than he is in *Boulder 1*. In *Boulder 2*, on the other hand, what *A* does that is independently wrong is to lift *B* up and move her without her permission. Lifting her up and moving her

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

is not merely a side effect of his shifting the danger away from himself; it is actually the means by which he does so.

The distinction between means and side effects has proven difficult to clarify in moral philosophy, especially for purposes of formulating the dubious doctrine of double effect. For our purposes, however, this distinction can be made clear enough as follows. To determine whether what makes A's act independently wrong counts as a means or a side effect, we are to ask ourselves this question: Would what *A* does to shift the danger still have successfully shifted it if, hypothetically, what makes this act independently wrong, and that only, were missing? The only qualification here is this. In asking ourselves this question – that is, in asking whether what *A* does to shift the danger would still have shifted it successfully if what makes this act independently wrong were missing – we may not, hypothetically, 'eliminate' what makes the act independently wrong simply by supposing that, being fully informed, *B consented* to it. The very existence of the ducking puzzle presupposes the absence of any such consent. If the answer to the question of whether what *A* does to shift the danger would still have shifted it successfully if what makes this act independently wrong, and that only, were missing is 'Yes', then what makes this act independently wrong counts as a mere *side effect* of his shifting the danger. If the answer is 'No', then it counts as the *means* by which he shifts the danger. In *Boulder 2*, for example, what *A* does to shift the danger is to pick *B* up against her and move her in front of him, and what makes this act independently wrong is exactly the same thing – namely, that *A* picks *B* up against her will and moves her in front of him. Thus what *A* does to shift the danger would not have successfully shifted it if what makes this act independently wrong were missing. So, in this case, what *A* does that is independently wrong counts as the means by which he shifts the danger. On the other hand, in *Model Airplanes* what *A* does to shift the danger is to jump away, and what makes this act independently wrong is that it destroys *B*'s model airplane. Obviously, *A*'s jumping away would still have successfully shifted the danger even if his jumping away had not destroyed *B*'s model airplane. So, in this case, what makes the act independently wrong is a mere side effect of his shifting the danger.

The last piece to the solution – the final condition that must be met for *A* to be morally liable for any harm that *B* suffers as a result of *A*'s shifting a danger away from himself onto *B*, or an opportunity away from *B* onto himself – is this. Not only must the means by which *A* shifts the danger, or opportunity, be independently wrong, but this wrongness must be directed, in part at least, against *B* herself, not against someone else *only*. *A*'s independently wrong act must, in other words, violate the rights of *B*, either of *B* exclusively, or of *B* and others. This condition

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

fails to be met in the following case; call it *Boulder with Threat*. *Boulder with Threat* is the same as *Boulder 1*, except that when *A* attempts to jump away from the oncoming boulder, he discovers that his foot is firmly caught in protruding roots, preventing him from getting away by himself. So *A* quickly whips out his pistol and threatens to kill *T*, an innocent bystander, unless *T* immediately pulls *A* free of the roots and away from the boulder's path, which *T* does. The boulder then hits and kills *B*, the next person in line. Assume, for the sake of argument again, that, in this case, *A's* forcing *T*, at gunpoint, to pull him away from the boulder's path is morally wrong. But, considered apart from any resulting harm to *B*, this wrong, this *independent* wrong, does not violate the rights of *B*. If it violates anyone's rights, it violates the rights of *T*, a third party in no danger at all from the boulder. And if it did violate *T's* rights, *A* might deserve to be reprimanded somehow for this violation. But, as confirmed by our moral intuitions, *A* does not deserve to be held morally liable for *B*'s death. *A* may be guilty of wrongfully pulling a gun upon *T*, but *A* is no more guilty of wrongfully killing *B* than he is in *Boulder 1*.

With the last missing pieces to the ducking puzzle now in hand, the solution proposed here is as follows. If person *A* causes person *B* to suffer harm either from a danger that *A* shifts away from himself onto *B*, or from losing an opportunity that *A* shifts away from *B* onto himself, then *A* is morally liable for the harm if and only if, in shifting the danger, or opportunity, *A* performs an independently wrong act against *B*, and what makes this act independently wrong is the means by which *A* shifts the danger or opportunity. Although this solution focuses only on morality, essentially the same solution is probably applicable to the same puzzle in law.

IV. THE EXCEPTION FOR FAIR COMPETITION

The rationale behind this solution to the ducking puzzle is as follows. Generally speaking, we are morally obligated not to harm others, not to damage their bodies, property, or financial position. But there are exceptions. Among them are an exception for self-defense, an exception for justified punishment, and an exception for permission. What is claimed here is that this solution to the ducking puzzle uncovers still another exception. It is an exception applicable to certain cases of shifting a danger away from oneself onto another, or an opportunity away from another onto oneself. A tragedy of human existence is that, all too often, it is a zero-sum game, a game in which one 'competitor' can benefit only by harming another. For example, in *Boulder 1*, by jumping to safety *A* benefits himself only by harming *B*, who, as a result of *A*'s jumping to safety, loses her life. And in *Widget 1*, by making a better

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

widget *A* benefits himself only by harming *B*, who, as a result, loses a considerable amount of business. For morality to prohibit these ways of benefitting oneself would be to prohibit competition that is a natural, and probably necessary, part of everyday life. Perhaps even more to the point, this competition can also be enormously valuable socially, as in the case of the economic competition exemplified by *Widget 1*. So benefitting oneself at the expense of another, through competition, is not prohibited by morality, *but only if the competition is fair*. And, by competition that is 'fair', what is meant is competition where one person prevails over another by means that do not constitute an independently wrong act against the other. Provided only that the means are not independently wrong, a person, in competing to avoid a common danger, or take advantage of a common opportunity, is morally permitted to do what causes his or her 'competitor' any amount of harm, even death. This is a huge exception to our moral obligation not to harm others, an exception for fair competition.

But the exception for fair competition is subject to an important qualification, the need for which can be appreciated by considering the following case.

Firecracker 3: This case is the same as *Firecracker 1* except that, rather than *A*'s being in a highly crowded marketplace, *A* is in a field where the only other person nearby is *B* and, although *A* could easily have protected himself from the lighted firecracker that landed next to him by tossing it anywhere other than at *B*, he instead, out of sheer malice, deliberately tosses it directly at *B*, putting *B*'s eye out.

In *Firecracker 3*, *A* is no more responsible for originating the danger in question, a lighted firecracker, than he is in *Firecracker 1*, nor is what he does to shift the danger away from himself and onto *B* independently wrong any more than it is in *Firecracker 1*. Thus it would appear that what *A* does in *Firecracker 3* falls under the exception for fair competition just as it does in *Firecracker 1*. But, contrary to what is true of *Firecracker 1*, intuitively it appears that, in *Firecracker 3*, *A* is indeed morally liable for any harm to *B* that results from what he does to shift the danger. We can make sense of this intuition by appealing to the least-harm principle, explained in Section I. In *Firecracker 3*, *A* can avoid the danger to himself from the lighted firecracker by tossing it away either at *B* or somewhere other than at *B*. According to the least-harm principle, however, from among alternatives such as these, ones that *otherwise* are morally equal, morality requires that we choose the one that causes the least harm. Clearly the alternative that causes the least harm is tossing the firecracker somewhere other than at *B*; so this is the alternative that the least-harm principle requires that *A* choose. But since, in violation of the least-harm principle, *A* chooses to

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

toss the firecracker at *B* instead, *A* is morally liable for the resulting harm to *B*.

This qualification upon the exception for fair competition is similar to a qualification upon a much better-known exception to our obligation not to harm others – namely, the exception for self-defense. According to the exception for self-defense, one may harm a wrongdoer to defend oneself, or others, against the wrongdoing. But, just like the exception for fair competition, the exception for self-defense is qualified by the least-harm principle. If, for example, punching a thief attempting to steal your wallet will prevent the theft just as effectively as shooting him, the least-harm principle requires that you prevent it by punching him, the alternative that causes the least harm.

Boorse and Sorensen themselves briefly consider a solution to the ducking puzzle that is somewhat similar to the one presented here, but mention objections to it that they consider sufficient for rejecting it. Suppose, they say, *A* gets *B* to move into the path of the oncoming boulder to shield him not by lifting her up without permission and moving her into its path, but by luring her there through a 'greeting or a handshake', or a 'friendly arm around the shoulders'.[13] *A* would still be morally liable for the harm to *B*, but, it might be asked, would such means as these be independently wrong? The reply, briefly, is as follows. Handshakes, arms around shoulders, and the like, are supposed to be gestures indicating friendliness, or at least non-hostility. But deliberately luring another to her death hardly qualifies as friendliness. In doing so, *A* is only pretending to be friendly, or at least not hostile. By pretending to be what he is really not, *A* is deceiving *B*, and, with certain exceptions of course, deception is morally wrong; it violates a moral obligation. To violate this obligation, one need not actually *say* something that is false, all one need do is deliberately get another to *believe* something that is false. Here *B* is maneuvered by *A* into falsely believing that *A* is being friendly to her. Since, as presupposed here, the wrongness of deception does not necessarily depend upon anyone's being harmed by it, *A*'s deception is indeed independently wrong, which precludes the exception for fair competition from applying to these cases, thereby leaving *A* morally liable for *B*'s death after all. A number of objections such as this one have been addressed in detail elsewhere.[14] Here, however, rather than addressing these objections yet again, let us, for the sake of argument, assume they have been addressed successfully, and proceed. What shall now be argued is that the same theoretical framework put to use here for solving the ducking

[13] Boorse and Sorensen, 'Ducking Harm', p. 124.
[14] Haslett, 'Murder and the Exception for Fair Competition', sects. 3–7.

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

puzzle can be put to use for solving Thomson's version of the trolley problem as well.

V. A SOLUTION TO THE TROLLEY PROBLEM

Consider the following case.

Bounder 3: A boulder is hurtling toward *A* who, with his foot hopelessly caught in protruding roots, is unable to move away. *A* realizes that the only way he can prevent himself from being killed by the boulder is by quickly moving a large rock in front of himself that will cause the boulder's trajectory to change just enough so that it will miss him. But *A* also realizes that, due to its change in trajectory, the boulder will inevitably hit and kill another person, *B*, whom it otherwise would have missed. *A* nevertheless moves the rock in front of himself, thereby saving his own life, but at the unavoidable cost of causing the boulder to hit and kill *B*.

This case should sound familiar; it is, in every morally relevant respect, the same case as *Firecracker 1*, a case in which *A*, unable to move himself away from the danger, instead moves the danger away from himself. Just as in *Firecracker 1*, *A*, in *Boulder 3*, moves the danger away from himself by means that are not independently wrong; there is nothing independently wrong about *A*'s simply moving a rock in front of himself. Thus, just as in *Firecracker 1*, in *Boulder 3* the exception for fair competition successfully precludes *A* from being morally liable for *B*'s death.

Notice, incidentally, that the exception for fair competition still applies even if, by shifting the danger, *A* necessarily would cause not just *B* to be killed, but four other people next to *B* as well (*Boulder with Five Killed*). Shifting the danger in this case would not be saintly to say the least, but, since *A*'s means of doing so would be no more independently wrong than in the case where only *B* would be killed, and since no alternative existed by means of which *A* could have shifted the danger by doing less harm, the exception for fair competition still would successfully preclude *A* from being morally liable.

Consider now a case exactly like *Boulder 3* except that, instead of saving just himself by shifting the danger onto one other person, *B*, *A* also saves four others next to him. Call this case, *Boulder with Five Saved*. Naturally the exception for fair competition applies here as well. Moreover, this case gets us closer to *Bystander at the Switch*, the very case that gives rise to Thomson's version of the trolley problem. A case that gets us still closer is the following.

Stuck Near Switch: This case is the same as *Bystander at the Switch* except that *A* is not a bystander, but one of the five people stuck on the main track, and he throws the switch himself by reaching out from where he is stuck.

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

The exception for fair competition applies, of course, to *Stuck Near Switch*, just as it does to *Boulder with Five Saved*, relieving *A* of any moral liability for the resulting death.

But now we come to the crucial question: Does this exception apply in *Bystander at the Switch* as well? Does this exception allow a bystander, on behalf of people who are facing a danger, to do whatever these people would themselves be morally permitted to do to avoid the danger? Keep in mind here that, as we saw near the end of the last section, the exception for fair competition is analogous in certain respects to the exception for self-defense. And it is well established that the exception for self-defense does allow a bystander, on behalf of people in jeopardy of being wrongfully harmed, to do whatever these people would themselves be morally permitted to do to defend themselves against the harm. For morality to permit bystanders, in defense of others, to do whatever these others would be permitted do in defense of themselves makes sense because, as a result, less harm, overall, from *wrongdoing* is likely to occur. Consider now the exception for fair competition. Because, as a result, less harm, overall, from *dangers* is likely to occur, it likewise makes sense for morality to permit bystanders to shift dangers away from people to whatever extent these people would themselves be permitted to do so by the exception for fair competition, but with one important limitation. As we have seen, given that the constraints of fair competition are met, people who themselves are facing dangers are normally permitted to shift them even if, by doing so, more harm is likely to occur as a result. Since, however, bystanders are neutral observers, they should be limited by morality to shifting dangers only if, by doing so, *less* harm is likely to occur as a result. Assuming other things were equal, this limitation would mean that bystanders would be permitted to shift life-threatening dangers away from more people onto fewer, but would not be permitted to shift them from fewer people onto more, or even onto the same number of people.

But what if other things are not equal? What if, for example, the one stuck on the spur track is a brilliant scientist about to discover a cure for cancer, while the five stuck on the main track are just ordinary people (*Scientist*)? Or what if the one stuck on the spur track is an ordinary person, while the five stuck on the main track are heinous members of the Mafia who, if not killed by the trolley, are likely continue murdering innocent people (*Mobsters*)? Would it, in these cases, count as 'less harm' to save the five at the cost of killing the one? Cases such as these indicate that, although bystanders should be limited to shifting dangers only if, by their doing so, less harm is likely to occur as a result, what exactly should count as 'less harm' here needs clarification.

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

To clarify this, we must first distinguish between being harmed by a danger 'directly', and being harmed by it 'indirectly', a distinction that the following example illustrates. Suppose a boulder is tumbling down a mountainside where a woman, Jane, is standing in its path. Jane is the devoted mother of three children, all of whom are safely at home. If the boulder hits and kills Jane, it harms her directly. Jane's children are then harmed also, since they will be without a mother. But Jane's children are harmed indirectly. To what extent should indirect harm be taken into account in determining what will result in less harm?

So far the argument has been this. Provided the exception for fair competition applies, people facing a danger are morally permitted to shift it onto others. And a bystander is morally permitted to shift a danger as well, but only if two conditions are met: (1) those actually facing the danger would themselves be morally permitted to shift it; and (2) by shifting the danger, less harm is likely to occur. In determining whether this second condition is met, thus making it *permissible* for a bystander to shift a danger, only direct harm should, it is suggested here, be taken into account. But even if it is morally permissible for a bystander to shift a danger, it will still not be morally *obligatory* that he do so.[15] Therefore, although morally permitted to shift the danger, the bystander will also be morally permitted not to shift it. And it is here – that is, for determining whether to choose not to shift it – that *indirect* harm should, it is suggested, be taken into account. In cases such as *Scientist* and *Mobsters* – cases in which it is fairly clear that the *indirect* harm resulting from shifting a danger outweighs the *direct* harm that shifting it prevents – a bystander, even if morally permitted to do so, should probably choose not to shift the danger. And if, in such a case, the bystander nevertheless does shift the danger, then, even though he may not have violated any moral obligation, he will probably be guilty of, and blameworthy for, having exemplified bad values.[16] We need not, however, pursue questions about indirect harms further here. Indirect harms are of more theoretical than practical interest; rarely will a bystander, with perhaps only seconds to decide whether to shift a danger, have enough information about indirect harms to override the obvious direct harms at issue.

[15] For an explanation of the difference between moral 'obligations' and moral 'values', see Haslett, 'Values', esp. sect. 3.

[16] If the person standing by the switch had a special obligation of care to one of the five, a 'commitment' obligation – if, say, one of the five were the person's eight-year-old daughter – then this person would indeed be obligated to shift the danger. On the other hand, if the *one* were the person's eight-year-old daughter, then this person would be obligated not to shift it. But, as we have seen, with a special obligation such as this to one of those stuck on the track, this person would no longer qualify as a 'bystander'.

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

A theoretical framework sufficient for solving Thomson's version of the trolley problem is now in place. At its core is the exception for fair competition. According to this exception, in *Stuck near the Switch* it would be morally permissible for any one of the five people stuck on the main track to save himself and the four others by reaching up and throwing the switch, thereby killing the person stuck on the spur track. And since, in *Bystander at the Switch*, the bystander's throwing the switch results in less direct harm than not doing so, throwing the switch is morally permissible for the bystander as well.

Of course, in *Transplant*, for Allie to kill Bailey to save the five patients also results in less direct harm than not doing so. But killing Bailey is not morally permissible. To see why, consider yet another case.

Transplant 2: This case is just like *Transplant* except that the person who shoots and kills Bailey so that his organs can then be transplanted into the five patients is not Allie, a bystander, but Alex, who is himself one of the five patients.

Transplant 2 is similar in one respect to *Stuck Near Switch*; in both cases, the person who causes the harm in question is one of the very people whose life is thereby saved. But *A*, in *Stuck Near Switch*, can avoid moral liability for the harm by appealing to the exception for fair competition, while Alex, in *Transplant 2*, cannot. The exception for fair competition relieves an agent from liability for harm he causes only if, without doing anything independently wrong, the agent causes the harm by shifting an existing danger away from some onto others, a danger the agent is not responsible for having originated. In *Transplant 2*, there is indeed an existing danger not originated by Alex that he *removes* from himself and the other patients; namely, the danger from diseased organs. But the harm to Bailey that Alex causes does not come from shifting this existing danger onto Bailey, as would be necessary if the exception for fair competition were to apply. It comes instead from the bullet that Alex fires at Bailey, a serious danger that Alex does not shift, but deliberately originates. Therefore Alex cannot avoid liability for this harm by appealing to the exception for fair competition or to anything else.

Since Alex, in *Transplant 2*, cannot avoid liability by appealing to the exception for fair competition or to anything else, Allie, in *Transplant*, who is a mere bystander, cannot, of course, either. Yet, for the reasons we have seen, the bystander, in *Bystander at the Switch*, can indeed avoid liability by appealing to the exception for fair competition. This

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

then is the relevant difference between *Transplant* and *Bystander at the Switch*.[17]

It may be useful to test the fair-competition solution set out here to the trolley problem by seeing how it works for another trolley case often brought up in discussions of this problem – namely, the case of *Fat Man*. In this case, a bystander, *A*, spots a fat man who happens to be standing nearby, and quickly pushes him onto the track in front of the runaway trolley, knowing that he is heavy enough to stop the trolley before it reaches the five, saving them, but killing him. Most people's moral intuitions tell them that *A* acted wrongly. But can the theoretical foundation developed here for explaining people's intuitions about the ducking puzzle and Thomson's version of the trolley problem also explain people's intuitions about *Fat Man*? To answer this, consider first a variation of *Fat Man* in which it is not a bystander who stops the trolley by pushing the fat man onto the track, but is instead Amos, one of the five stuck on the track, who, by reaching out and grabbing the fat man, pulls him onto the track. In this case – call it *Stuck Near Fat Man* – Amos, as opposed to Alex in *Transplant 2*, does indeed shift a danger away from himself and onto the victim, an existing danger he did not originate. Amos, nevertheless, is just as morally liable for the victim's death as Alex is in *Transplant 2*. This is because *Stuck Near Fat Man* is, in all morally relevant respects, exactly like *Boulder 2*. Just like the agent in *Boulder 2*, what Amos does to shift the danger away from himself is to move the victim, against his will, into the path of the danger so that the victim shields him from it. But moving a person, against the person's will, violates a moral obligation *by itself*; that is, even if it does not result in any harm. Thus, exactly like the agent in *Boulder 2*, what Amos does to shift the danger is *independently* wrong. And since it is independently wrong, this precludes the exception for fair competition from applying, thereby leaving Amos morally liable for any foreseeable harm from what he does. Amos, in *Stuck Near Fat Man*, is therefore morally liable for the death of the fat man, and, a fortiori, so is the bystander, *A*, in *Fat Man*. The theoretical foundation developed

[17] Thomson herself, in 'The Trolley Problem', appears to have been working toward a solution somewhat like the one defended here. See especially p. 285. She now claims, however, that the trolley problem is not really a problem after all. (See Judith Jarvis Thomson, 'Turning the Trolley', *Philosophy & Public Affairs* 36 (2008), pp. 359–74.) What she now claims is that, in spite of most people's moral intuitions to the contrary, it actually is no more permissible for *A*, the bystander in *Bystander by the Switch*, to kill the one to save the five stuck on the track than it is for Allie, the bystander in *Transplant*, to kill Bailey to save the five in need of transplants. This is because, she says, you should never do a good deed (such as save the five stuck on the track) unless you bear the entire 'cost' yourself. But, if the theoretical framework set out here is sound, cases such as *Bystander at the Switch* are an exception to this otherwise sound generalization, an exception derivable, ultimately, from the exception for fair competition.

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press

originally for explaining our moral intuitions about the ducking puzzle thus succeeds in explaining our moral intuitions about these other trolley cases as well.

Finally, if the very same theoretical framework developed for explaining our moral intuitions about the ducking puzzle also succeeds in explaining our moral intuitions about trolley cases, this reinforces its original application to the ducking puzzle. The more such intuitions a theoretical foundation succeeds in explaining, the more likely it is that the foundation is sound.

dhaslett@udel.edu

https://doi.org/10.1017/S0953820811000148 Published online by Cambridge University Press