

Reflective Moral Equilibrium and Psychological Theory*

Mark van Roojen

Tamara Horowitz discusses and criticizes Warren Quinn's use of thought experiments to support a version of the Doctrine of Doing and Allowing.¹ She argues for the empirical explanatory hypothesis that another set of principles explains our common agreement on the correct outcome in those thought experiments. And she remarks along the way that therefore, contrary to Quinn's own conclusions, our intuitions concerning those examples do not provide support for the Doctrine of Doing and Allowing. In the course of the paper, Horowitz rebuts an objection from an anonymous reader that her argument misconstrues the role of such examples in the process of reaching a set of conclusions in reflective equilibrium.

I believe that something like the reader's objection is correct. In support of the objection I will argue the following: when we use the method of reflective equilibrium to reason to moral principles, thought experiments such as those discussed by Horowitz function primarily to yield first-order moral judgments that actions of this or that type are right or wrong. These judgments do not require or presuppose the truth of psychological or explanatory claims about how we come to accept those first-order judgments. We then look for principles that would explain their truth as opposed to explaining our acceptance of them. Thus psychological explanations do not initially play a role in justifying normative conclusions. However, such explanations of our particular judgments can be relevant if they show that our reaching the initial judgments involved some sort of error. But showing that we are making such errors will require more than a psychological explanation of our beliefs, even one which does not rely on the moral principle we have used the particu-

* I would like to thank Robert Audi, Albert Casullo, Harry Ide, and Joe Mendola for helpful comments on earlier drafts of this article. I would also like to thank the editor of *Ethics* and an anonymous reader for their patient comments on several drafts which served to improve the final article considerably. Finally, I would like to thank Gilbert Harman for discussion of some of the implications for ethics of prospect-theoretic explanations of moral judgments.

1. Tamara Horowitz, "Philosophical Intuitions and Psychological Theory," *Ethics* 108 (1998): 367–85. Page references in the text are to this paper.

lar judgments to support. A case will need to be made that the reasoning postulated by the psychological explanation is inherently fallacious.

Insofar as Horowitz offers no argument that prospect-theoretic reasoning is fallacious, she has not made a case for undermining Quinn's main conclusions. Furthermore, when we look at prospect theory for ourselves, we should conclude that such reasoning as it predicts is not inherently fallacious where we can make distinctions of the sort the doing/allowing distinction provides. Thus objections to prospect-theoretic framing of other sorts of choices do not extend to the examples used by Quinn. Drawing normative moral conclusions from the examples as Quinn does appears to be justified even if Horowitz's psychological explanatory claims are true.

Quinn suggested that reflection on two thought experiments, one involving killing someone to save several others and the other involving letting one die to save the same number of others, supports the idea that there is a morally significant difference between bringing about a death oneself and merely allowing it to happen.² The examples in the thought experiments support this idea because, as reflection shows, it is wrong to kill the one to save the others but not wrong to let the one die to save them. After canvassing several other similar examples, Quinn concluded that a principle which best captures and explains this difference is a version of the Doctrine of Doing and Allowing. This doctrine postulates constraints on "positive" agency which brings about a result even where there is no constraint on negative agency which brings about a similar result. Killing is a paradigm case of such positive agency whereas letting die only involves negative agency. After qualifying and amending the proposal a bit, Quinn completed his argument by offering a more philosophical rationale for the principle invoked by the doctrine.

Horowitz's main thesis in the paper may be explanatory. But she also objects to Quinn's claim that verdicts about what to do in the scenarios introduced in the thought experiments support the Doctrine of Doing and Allowing. She argues that an alternative more general and hence preferable explanation of his responses to the scenarios is available, and that this explanation will undermine his endorsement of the principle. The general explanation by virtue of invoking prospect theory³ is a psychological explanation of Quinn's responses to the examples. The key prospect-theoretic idea is that people tend to evaluate outcomes from

2. Quinn's argument can be found in "Actions, Intentions, Consequences: The Doctrine of Doing and Allowing," *Philosophical Review* 98 (1989), reprinted in his *Morality and Action* (Cambridge: Cambridge University Press, 1993), pp. 149–74. While I am defending Quinn's use of thought experiments in his argument, I don't wish to take a position on his overall interpretation of the Doctrine of Doing and Allowing.

3. Prospect theory is a nonstandard sort of decision theory defended by Daniel Kahneman and Amos Tversky in "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47 (1979): 263–91.

some sort of neutral baseline and that positive deviations from that baseline are regarded as less significant than negative deviations. Horowitz uses this idea to explain Quinn's verdicts regarding the thought experiments in the following way: on the assumption that the overall verdict is a function of the values of components of the overall outcomes, Horowitz divides the problem up into a component which involves comparing the value of killing one person as opposed to sparing him with the value of saving one person as opposed to letting him die and another component which is supposed to be present in both cases.⁴ Then she offers the hypothesis that each comparison highlights a different baseline from which to measure gains and losses. Killing is judged relative to a baseline in which the person is alive, saving relative to one in which the person is dead. Since gains are judged by most people to be less significant than losses, prospect theory predicts that killing will count more negatively against a course of action than merely letting someone die.

From here, Horowitz's reasoning seems to be that prospect theory should displace the Doctrine of Doing and Allowing as the best empirical explanation of people's judgments with respect to such examples because it is more general, covering a wider variety of examples than just those employed by Quinn, and because it is well confirmed as a set of principles that do as a matter of empirical fact govern people's reasoning in those cases. "By the ordinary epistemic standards of decision psychology, the prospect-theoretic explanation is the one we should accept" (p. 380). So far her point is about psychology and psychological explanation.

But she clearly wants to make a broader normative point. In particular, she argues that given the psychological claim that prospect theory explains our intuitions in Quinn's examples, the examples do not support the Doctrine of Doing and Allowing at all.

It is crucial for Quinn that the intuitions he elicits be moral intuitions, since he wants to argue that our moral intuitions both support the Doctrine of Doing and Allowing and conflict with consequentialism. Further, he wants the doctrine of rights which he develops in his paper to explain these very moral intuitions. But to the extent that the intuitions elicited by Quinn's thought experiments are explained by prospect theory, they are not moral intuitions at all. (p. 381)

4. I actually have quite a bit of trouble figuring out from her text exactly what the second component is supposed to be and what she has to say about it. I believe her idea is that one option in each choice involves saving five people, while the other option in each choice involves failing to save five people. These features make up the second component which the examples have in common. In any case, the text is clear that the first component involves a comparison of the value of a death due to killing as opposed to a death due to letting die. And that is really the component doing the explanatory work on the prospect-theoretical account.

And she responds to a reader's objection that despite her psychological claims, "these differing responses to the Rescue Dilemmas can still play a role in ethical argumentation if the argumentative method being used is one of reflective equilibrium," with the claim: "My contention is that when Quinn, or anyone else, judges that there is a difference in what it is permissible to do in the two Rescue Dilemmas, they are mistaken in thinking that they are making a moral judgement at all" (p. 381).

This passage is a bit puzzling partly because the distinction between moral intuitions and nonmoral intuitions is not all that clear. My interpretation is that Horowitz means to contrast intuitions grounded in a moral principle such as the Doctrine of Doing and Allowing with intuitions grounded in principles such as those used in prospect theory.⁵ Her idea then would be that intuitions caused by prospect-theoretic reasoning cannot play the role that moral intuitions normally play in reaching reflective equilibrium on moral matters. Thus she thinks that her argument will undermine not only Quinn's particular argument based on these examples, but any use of such examples to vindicate the Doctrine of Doing and Allowing.⁶ Her reasoning here is not entirely idiosyncratic.

5. A slightly different interpretation would contrast judgments grounded in one sort of psychological mechanism or competence with those grounded in another. Someone might think that certain moral judgments resulted from mechanisms that are universal, much as some linguists postulate universal linguistic mechanisms to explain linguistic judgments. On such a view, such moral mechanisms or competences could interact with other components of the person's psychology to generate moral judgments in particular cases. If this is the idea, Horowitz might be distinguishing judgments depending on such subpersonal mechanisms or competences from those that depend on other mechanisms. On this reading, Horowitz would be claiming that the judgments do not stem from specifically moral competences or mechanisms. My arguments in the text would have the same upshot against this way of making the distinction as well. For more on subpersonal moral competences and parallels with linguistics, see Gilbert Harman, "Moral Philosophy and Linguistics," forthcoming.

6. Someone might think that I am being uncharitable in attributing to her a more ambitious argument which attempts to undermine drawing normative conclusions from thought experiments like Quinn's. Rather, the objection would go, all she is doing is offering prospect theory as a psychological explanatory theory which displaces the Doctrine of Doing and Allowing as a competing psychological explanation. Against this I have several replies: first, on my reading of her argument, Horowitz makes both a more modest empirical explanatory claim and the more ambitious claim that the modest claim undermines any attempt to reach normative conclusions from such thought experiments. Hence, textual evidence that she makes the modest claim is not in conflict with my interpretation. Second, a purely modest reading of her argument ignores claims she does make, such as when she replies to the anonymous objector in the way I quote above. Whatever the point there is, it is aimed at *anyone* who tries to argue for moral principles from examples like Quinn's. And if Horowitz's ambitions were not to undermine normative claims reached through such reasoning, the better response to the objection regarding reflective equilibrium would just be that that process is irrelevant to her argument since its conclusions are normative while hers are psychological. Further, such textual evidence comes from her concluding paragraphs: "It is therefore possible that prospect theory . . . provides the correct account of the reasoning engaged in by people who come to have Quinn's intuitions concerning his Res-

Robert Nozick, whom Horowitz cites, includes an aside in his own discussion of constraints on harming that suggests that if they are prospect theory framing effects he would no longer endorse them.⁷

It is here that I want to focus my disagreement: prospect theory does not undermine arguments from our conclusions about what is right in the thought experiments to the Doctrine of Doing and Allowing. My point is not that prospect theory cannot tell us something about our thinking regarding the distinction marked in the examples, or even that it might not be used to debunk the moral import of the distinction. But, I will argue, the mere fact that prospect theory can explain our having these responses to hypothetical examples does nothing to undermine the use of such examples in an argument for that distinction. Prospect theory is not an explanatory rival to the Doctrine of Doing and Allowing as a moral principle precisely because it explains, if it does, as part of an empirical psychological theory and not as part of a normative justificatory theory.

There are two senses in which a moral principle can explain our intuitions regarding moral thought experiments. A principle might explain the psychological fact that we make certain judgments. Or a principle might explain the truth (if it be such) of the content of the judgment.⁸ We might be interested in explaining the fact that we tend to think that running over one to save five is wrong or we might rather be interested in explaining the fact that running over one to save five is wrong. When we try to reach reflective equilibrium between our moral principles and our specific intuitive moral judgments, explanations of the second sort are especially what we are looking for.⁹ One problem with Horowitz's discussion of Quinn's argument is that it leaves out the role of judgments with purely moral content (e.g., judgments that a particular action is right or wrong) in reaching conclusions about moral

cue Dilemmas. If this is so, then Quinn's philosophical thought experiments do not provide us with an argument for his philosophical conclusions" (p. 385). Given that Quinn's philosophical conclusions were normative, the most charitable reading of Horowitz's claims is one in which she also means to take issue with his normative arguments.

7. Robert Nozick, *The Nature of Rationality* (Princeton, N.J.: Princeton University Press, 1993), p. 60.

8. Gilbert Harman in *The Nature of Morality* (New York: Oxford University Press, 1977) notes the same sort of distinction between facts to be explained when we are talking about moral observations and their explanation: "It is . . . important to note an ambiguity in the word 'observation.' You see the children set the cat on fire and immediately think, 'That's wrong.' In one sense, your observation is that what the children are doing is wrong. In another sense, your observation is your thinking that thought. Moral principles might explain observations in the first sense but not in the second sense" (p. 8). I thank the editor of *Ethics* for reminding me of this passage.

9. I suspect that the anonymous reader cited by Horowitz as complaining that her argument presupposes a mistaken view of reflective equilibrium aimed to make something like this point.

principles. Instead it focuses solely on higher-order claims about what people think, claims which attribute thoughts with those moral contents to people.¹⁰ Without careful argument to show that psychological claims about their causation undermine our right to the first-order moral claims, there can be no showing that using such judgments to justify moral principles is out of order.

Prospect theory taken as a psychological theory (as Horowitz intends it) might edge out another theory aimed at explaining the fact that we have the intuitions that we do.¹¹ But it does not automatically and directly displace the Doctrine of Doing and Allowing because that doctrine belongs to a normative theory, which is aimed at explaining the truth of the first-order moral content. If the explanandum were a certain psychological claim about what a person or some people believe, then we would be looking for a psychological or causal explanation of how they came to believe that. On the reasonable assumption that a more general psychological theory well confirmed in other areas is preferable to a less general theory attempting to explain the same thing, the prospect-theoretical explanation might well emerge as the best theory and hence perhaps the only theory we should accept as the best explanation of the empirical phenomenon to be explained.

10. Am I denying that Quinn's argument depends on the "assumption that people who share his intuitions in the case of Rescue Dilemma 1 and Rescue Dilemma 2 do so because they accept, however inexplicably, the Doctrine of Doing and Allowing" (Horowitz, p. 369)? Well, yes and no. Quinn is doing two things in the first section of his paper. He was clarifying the content of a doctrine that he thought was widely accepted, and he was arguing for it. Insofar as he was clarifying what he thought was a widely held set of views, he assumed that his reactions to the cases were expressions of those views. But in his argument for the doctrine from those dilemmas, the assumption does not play a role. At least this seems to me the best reading of what is going on. It is somewhat understandable that Horowitz reads him differently. Quinn's own discussion seems to me to run the two roles together somewhat. He sometimes speaks in the same breath of what we may do and of what we will tolerate someone doing (see Quinn, p. 365, for example). But even if Quinn himself does not stay sufficiently clear on the point, insofar as Horowitz wants her argument to apply to any argument from the same examples to similar conclusions, we should be careful to keep the two projects apart.

11. Whether it does or not is actually also a tricky question, even should prospect theory offer a correct psychological explanation of our moral judgments over a particular range of cases. This is because it is not obvious what it amounts to for a certain principle or set of principles to control our intuitive judgments when those judgments do not result from conscious consultation of the principle. Depending on what this sort of explanation entails, it could be compatible with equally correct explanations employing different principles. For example, if the claim is cashed out in terms of the response depending counterfactually on the judgment rendered agreeing with the principles employed in prospect theory, principles which necessarily lead to the same choices will offer equally good explanations. If the claim is that the biological process which realizes the thought process has a structure that mirrors the structure of the reasoning employing the principle, it would be possible for a structure to be isomorphic to the structure of different explanations invoking different principles.

But if what is to be explained is the propositions we believe, *that saving five by killing one is wrong but that neglecting to save one in order to save five is not wrong*, then prospect theory has not done the job. Or, at least, the psychological theory that people do reason as prospect theory describes does not explain these propositions. Because it does not explain these propositions it is not in direct competition with the Doctrine of Doing and Allowing, even when the latter is proposed as the best explanation of the moral propositions belief in which the prospect-theoretical account purports to explain. A prospect-theoretic normative theory which conflicted with the doctrine would be in direct competition with it. But a normative version of prospect theory is not obviously something that could be confirmed purely by the empirical methods Horowitz champions, nor does she present prospect theory as such a normative theory.

The general point is that reflective equilibrium aimed at moral conclusions starts from the content of our belief, in this case that killing five to save one is wrong, and works toward principles that fit with those contents. From there the process normally involves looking for principles that explain those contents themselves, not the fact that we believe them. Claims about what we believe or about why we believe it do not contradict the judgments we start from, nor do they contradict the claim that this or that principle explains why those judgments are true. Thus, they do not directly falsify any of the judgments we normally make in reaching reflective equilibrium.

This is not to say that higher-order reflection on first-order judgments can play no role in deliberating about what to accept. It can undermine our conclusions in a more indirect way. Among the ways that we might come to doubt the content of a judgment is to note that we made it in conditions that were not apt for judging rightly. But not every causal/psychological explanation of our first-order judgments will justify such debunking.¹² We will need some reason to think that the particular explanation in question shows our judgments to be distorted or based on illusions.

Let's examine whether prospect theory provides the grounds for an argument of this sort. There are two possibilities to consider. Either the judgments predicted by prospect theory conflict with those yielded by the Doctrine of Doing and Allowing for some range of cases or they do not.¹³ If they conflict, and if we had an argument for prospect-theoretic

12. For an interesting discussion of debunking explanations, see S. L. Hurley, *Natural Reasons: Personality and Polity* (Oxford: Oxford University Press, 1989), pp. 288–313.

13. You might think that Horowitz rules this possibility out insofar as she is trying to explain our responses to the same cases that Quinn and others use the Doctrine of Doing and Allowing to justify. But she explicitly notes that she claims only to explain the cases that Quinn himself employs to support the doctrine, and leaves open the possibility that other doing/allowing verdicts would not be similarly explicable. On this see the quotation in the next footnote.

principles as correct principles of practical reasoning, we might have reason to abandon the Doctrine of Doing and Allowing. But Horowitz does not argue that there is any such conflict and nothing in her article suggests that she offers prospect theory as a correct normative theory of practical reasoning. So I won't pursue the possibility further beyond remarking that arguments to vindicate prospect theory as a normative theory will likely have a lot in common with arguments for the Doctrine of Doing and Allowing.

Suppose then that they do not conflict. How would this undermine our normative conclusions? Perhaps prospect-theoretic reasoning is inherently fallacious, though Horowitz provides no argument that it is. Whether that should undermine a proponent's commitment to the Doctrine of Doing and Allowing depends on a number of issues. Is the only reason we are inclined to treat killing and letting die differently the fallaciousness of prospect-theoretic reasoning? Or, can we see independent reasons for treating killing differently from letting die? It is often possible to realize that one is reasoning in a fallacious way, eliminate the influence of the fallacy on one's reasoning, and find an independent argument for reaching the conclusion in question from the same evidence. Someone tells me I only favor a certain policy because it favors the poor, and I am poor. I know that self-interest often has this sort of influence on our thinking, and I agree that it is irrelevant. Yet upon reflection I think that the policy's favoring the poor is a good reason to favor it nonetheless because fairness also requires favoring them. Similarly even if prospect theory explains the fact that we have Quinn's intuitions and even if prospect-theoretic judgments are generally fallacious, we might be rational to accept and reason from the intuitions we find intuitively plausible nonetheless.

This possibility for the case at hand is not all that far-fetched since Horowitz does not claim that prospect theory explains all the myriad other examples people have used to argue for the importance of a distinction between doing and allowing.¹⁴ Thus, if the advocates of the Doctrine of Doing and Allowing are correct, the doctrine yields verdicts that agree with many people's moral intuitions for a broad range of cases beyond those Horowitz claims to explain. If so, it has implications not just for our normative theorizing but also for the explanatory hypotheses we should accept. Responses to the cases Horowitz discusses might be causally overdetermined. There probably is a psychological explanation for the responses of people in all of the cases Horowitz does not purport to explain. If this explanation involves our relying on a principle like that advocated by fans of the Doctrine of Doing and Allowing and that principle also covers the cases Horowitz does claim to explain, then we might reasonably conclude that our responses to those latter cases were causally

14. "The requirement that prospect theory account for all of these cases would be question begging" (p. 380).

overdetermined. Supposing that to be so, and supposing that only one of the explanations involves tacit reliance on a fallacious principle of reasoning, we do not have sufficient reason to discard our intuitions about those cases as corrupt.

This all presupposes that prospect-theoretic reasoning is always fallacious. Horowitz doesn't actually argue that it is but it is instructive to briefly look at the issue for ourselves. We have some reason to think that *some* prospect-theoretic reasoning is fallacious when we look at Kahneman and Tversky's Asian disease experiment (discussed by Horowitz on p. 370), where people tend to treat the very same options differently depending solely on how they are described. Since the choices, however described, are the same, such responses are irrational.¹⁵

But we should take some care to notice that not every prospect-theoretic chain of reasoning is called into question by the example. As Horowitz tells us, prospect theory involves a multistage procedure, one of which includes framing the choice in such a way that certain outcomes are regarded as neutral. What seems to have gone wrong with the example here is that the very same option is framed in two inconsistent ways depending on how it is described. Nothing in the example shows anything wrong with treating losses from a neutral baseline differently from gains. Such reasoning might well be appropriate where framing proceeds in a reasonable manner. It is differential weighting that is common to the Asian disease example and the doing/allowing examples, not treating the same outcome in inconsistent ways, at least so long as outcomes can be individuated by features such as whether they involve my killing someone or not. And that they can is one of the things fans of the doing/allowing distinction wish to defend.¹⁶ If they are right, this sort of prospect-theoretic reasoning involves no fallacy. So the conclusion that prospect theory explains why we reason in this way does not entail that our reasoning is fallacious when we reach conclusions about the moral significance of killing and letting die.¹⁷

15. Our belief that it is, is not obviously an empirically justified belief, but rather a substantive and seemingly a priori claim about what sorts of differences could be rational to take into account when choosing courses of action.

16. All of this may justify a worry that prospect theory is not actually a rival theory at any level to theories which rely on a distinction between doing and allowing. Instead such theories might offer us a reason or an explanation of why we should or do frame choices involving killings differently from cases with similar outcomes that come about by our allowing deaths to occur.

17. It is important to see that my claim here is not just about the truth of the responses but about the validity or fallaciousness of the reasoning involved in reaching them. Even if incorrect ways of framing outcomes lead prospect-theoretic reasoning not to be generally reliable, the lack of reliability is due to the framing and not to the differential treatment of options once framing has occurred. Thus, where we have no reason to think that the framing of options involves a mistake, we have no reason to think that the reasoning is unreliable or invalid.

I have discussed this way of proceeding from Horowitz's empirical explanatory conclusions to her claim that examples like the rescue examples do not support the Doctrine of Doing and Allowing, even if it were reached through reflective equilibrium, because it seems to me the most plausible way of making such an argument. Horowitz, however, seems perhaps to have had another idea in mind when she claimed that her argument undermined using the examples to draw conclusions favorable to the Doctrine of Doing and Allowing. Rather than arguing that our reasoning about the examples is fallacious, she seems to want to claim that our judgments are irrelevant to the Doctrine of Doing and Allowing. In the passages I have quoted earlier she suggests that the prospect-theoretic explanation of people's reasoning about the rescue examples shows that they are not moral judgments or intuitions and hence not relevant to supporting a moral principle. The idea might be that some features of a set of options might be moral features, whereas others are not. An intuition would be a moral intuition if it depended on a sensitivity to the moral features of an example, whereas it would not be if it instead depended on a sensitivity to features that would make a difference in choosing between options even where no moral issue is involved in the choice. Horowitz might be claiming that since our differential responses to the rescue cases depend on treating potential losses differently from foregoing a potential gain, and since this contrast is present and influences choices between options without a moral difference such as that between killing and letting die, these responses are not moral intuitions.¹⁸

I find this strategy of argument less promising than a strategy that relies on showing the underlying judgments to be false. Relations of support or explanatory relevance between propositions need not respect boundaries between one class of choices which are classed as moral choices and other equally practical sets of choices which are in some sense nonmoral. So far as I can see, a *prima facie* relevant judgment should only be disregarded if it is undermined. Thus, if the judgments about particular examples are true, and if principles like the Doctrine of Doing and Allowing fit coherently with them and even entail them without contradicting other judgments which are also true, that doctrine can explain their truth. Reflective equilibrium may then reasonably be achieved by accepting the doctrine as the best explanation of what we believe to be true in the particular cases.

That the features to which we are sensitive in the rescue examples can be present in cases that do *not* involve choices between killing and letting die does not alter this. Supposing that they are present in such cases, one possibility is that they are also always present in relevant

18. I thank the editor of *Ethics* for suggesting that this sort of argument might have been what Horowitz intended.

thought experiments and explain all of our differential reactions to contrasting killing/letting die examples. There would then be a necessary connection between being an instance of either killing or letting die, and possessing or lacking the relevant features. We might then see the Doctrine of Doing and Allowing as the application of some more general principle invoking those features to a more restricted domain of harming/not helping choice situations. That the doctrine and the judgments about the particular cases were both implications of this more general principle would neither undermine the doctrine, nor show the responses to the examples irrelevant to the truth of the doctrine. The overall explanation would still rely on the necessary connection between the moral features (that this is a killing, or that this is a letting die, for example) and the features cited in the more general principle and this probably would depend in part on the significance of just those moral features.

We can see this more concretely when we recall that prospect theory involves two components. One is framing a set of options to determine a baseline relative to which the options are compared, and the other involves differential treatment of gains and losses relative to a baseline arrived at in the framing stage. If Horowitz's prospect-theoretic explanation of responses to the rescue examples is correct, framing in those examples establishes different baselines depending on whether a death is caused by our killing a person, or by our failing to save her.¹⁹ So the prospect-theoretic explanation of our responses to the examples in question involves a sensitivity to "moral" features of the situation, and in fact just those features that Quinn's theory claims are relevant. If, as we are imagining here, a similar explanation could be offered for all of the relevant killing/letting die thought experiments, such moral features would still be part of our overall story for all of those cases.

Consistent with Horowitz's argument, however, is another possibility, that we cannot use prospect theory to explain what we believe to be correct responses to all of the relevant killing/letting die thought experiments. This is because the features employed in the prospect-theoretic explanation might only be present for some of the thought experiments that people typically use to support the Doctrine of Doing and Allowing.²⁰ Then there would be a couple of further possible options. If we also endorse judgments concordant with the Doctrine of Doing and Allowing where the features relevant for prospect theory are not present we might well have an example of explanatory overdetermination.²¹ If,

19. "In deciding whether to kill the person or leave the person alone, one thinks of the person's being alive as the status quo and chooses this as the neutral outcome. . . . But in deciding whether to save a person who would otherwise die, one thinks of the person's being dead as the status quo and chooses this as the neutral outcome" (pp. 377-78).

20. Again, this is a possibility that Horowitz specifically concedes as noted in n. 14.

21. Actually we should be a bit careful here. If my earlier argument is correct, there are two things we might explain, the truth of the contents of the judgments and the fact

on the contrary, we concluded that the judgment yielded by the doctrine were not correct for the cases where the prospect-theoretic elements were absent, it might make us hesitate to conclude that the doctrine is true based just on the examples highlighted by Quinn. But we should hesitate anyway in that case, even in the absence of prospect theory. For that would mean that we were disinclined in reflective equilibrium to endorse what seemed to be implications of the doctrine. Thus, I conclude that a strategy which uses prospect theory to show that our judgments about Quinn's examples are irrelevant to conclusions about the significance of doing and allowing is no more promising than one which tries to show that our judgments are fallacious by invoking prospect theory.

All of this leaves those who wish to use Quinn's examples to argue for the Doctrine of Doing and Allowing in a relatively secure position, Horowitz's remarks to the contrary notwithstanding. Challenging such arguments will require the challengers to make normative arguments of the sort that Horowitz does not pursue. Either they will have to make the arguments to vindicate prospect theory as a rival normative theory yielding conflicting actions as the right choices, or they can make the sorts of empirical explanatory arguments of the sort Horowitz makes to show that prospect theory explains judgments that killing is often wrong where letting die is not. But they must then supplement those arguments with normative arguments to show that such prospect-theoretic reasoning is inherently fallacious. Otherwise the particular intuitive judgments Quinn invokes to support the Doctrine of Doing and Allowing will remain as commitments to be put in reflective equilibrium with candidate moral principles that justify and explain those judgments.

that we make judgments with those contents. Explanation of either might be overdetermined, but for the explanation of the truth of the normative judgment to be overdetermined we would also have to accept the prospect-theoretic explanation as an example of good reasoning. We would likely need an argument that it is.